



The Value of Accuracy in Adjudication: An Economic Analysis

Citation

Louis Kaplow, The Value of Accuracy in Adjudication: An Economic Analysis, 23(1) J. Legal Stud. 307 (1994).

Published Version

doi:10.1086/467927

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10611788>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The Value of Accuracy in Adjudication: An Economic Analysis

Author(s): Louis Kaplow

Source: *The Journal of Legal Studies*, Vol. 23, No. 1, Economic Analysis of Civil Procedure (Jan., 1994), pp. 307-401

Published by: The University of Chicago Press

Stable URL: <http://www.jstor.org/stable/724326>

Accessed: 16/02/2010 19:33

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Legal Studies*.

THE VALUE OF ACCURACY IN ADJUDICATION: AN ECONOMIC ANALYSIS

LOUIS KAPLOW*

I. INTRODUCTION

CONCERNS for the accuracy of adjudication permeate analyses of procedural rules and aspects of substantive law. Yet the value of more accurate adjudication is largely taken for granted. When this is done, however, there is no basis for choosing among rules (or for judges to make discretionary judgments when applying them), for it typically is the case that greater accuracy comes at a cost. Even if precise quantification of various benefits of accuracy is impossible, decision-making will be enhanced by understanding why accuracy may be desirable.

This article presents an economic analysis of the value of accuracy in adjudication. It focuses primarily on three considerations: the effect of inaccuracy on implementation of substantive legal norms (individuals' incentives to comply with the law and proper assignment of future rights and obligations), the administrative costs of inaccuracy and attempts to reduce it, and costs arising from the imposition of sanctions.

Accuracy is a central concern with regard to a wide range of legal rules. One might go so far as to say that a large portion of the rules of

* Harvard Law School and the National Bureau of Economic Research. I am grateful for comments from Marcel Kahan, A. Mitchell Polinsky, Frederick Schauer, Steven Shavell, Kathryn Spier, and participants at the Conference, Harvard's Law and Economics Workshop, and the American Law and Economics Association Meeting; research assistance from Michael Steigmann; and support from the John M. Olin Foundation. Readers interested in the published comments by Daniel Ortiz and Stephen Bundy will find discussions of most of their points in Section IIA (especially note 9 and Subsections 1b, 1c, and 4) and, for the latter portion of Bundy's comment, Section VC (especially Subsection 3). Presented at the John M. Olin Program in Law and Economics Conference on "Economic Analysis of Civil Procedure" at the University of Virginia School of Law, March 26-27, 1993.

[*Journal of Legal Studies*, vol. XXIII (January 1994)]

© 1994 by The University of Chicago. All rights reserved. 0047-2530/94/2301-0014\$01.50

civil, criminal, and administrative procedure and rules of evidence involve an effort to strike a balance between accuracy and legal costs. The regulation of lawyers in litigation and some other settings is appropriately viewed, for present purposes, as an aspect of procedural rules concerned with achieving accurate outcomes while not incurring excessive costs. Implicit judgments concerning the value of accuracy are central in assessing major legal reforms (such as substituting an inquisitorial system for an adversarial system), evaluating more modest changes (limiting discovery or the use of expert witnesses), determining how adjudicators should exercise discretion (pretrial orders concerning the conduct of litigation), and designing and using alternative dispute resolution (often specified by contract).

Many aspects of substantive law also are concerned with accuracy. Most obvious are special burdens of proof in particular areas of law (such as burden shifting in employment discrimination cases or *res ipsa loquitur* in tort law) and rules concerning what evidence meets even conventional proof burdens (such as whether a dealer's complaint to a manufacturer is sufficient evidence of an antitrust conspiracy to reach a jury). Also important are components of substantive law that determine which categories of behavior are to be distinguished. For example, when assault with intent to commit murder is made a crime separate from simple assault, the legal system distinguishes two types of behavior, with the result that sanctions are more precisely tailored to individuals' conduct. When various categories of consequential damages are allowed, defendants' payments may reflect plaintiffs' actual harm more closely; but consequential damages sometimes are disallowed because they would be too costly to establish or may be so prone to error that defendants' payments would less accurately reflect differences in actual harm.

This article is primarily concerned with the question of why accuracy is valuable. It is assumed throughout that more accuracy can be obtained only at a higher cost.¹ No attempt is made to determine which legal rules are more accurate, by how much, and at what cost. Such inquiries are best made case by case. Rather, the analysis seeks to illuminate the following sort of inquiry: if a contemplated legal reform would increase accuracy in some specified manner and increase cost by a determined amount, is the reform desirable?

¹ Some changes in the legal system might make it simultaneously more accurate and cheaper, but it is usually obvious that such changes are desirable (from the economic perspective employed here) and thus analytically uninteresting to consider them.

Other closely related questions are considered as well. In particular, this article addresses how the cost of improving accuracy depends on other aspects of the legal system (such as the extent to which a law is enforced) and how problems of accuracy affect the appropriate design of other elements of the system (such as whether sanctions should be higher or lower when inaccuracy is a more serious problem). The analysis also examines expenditures on legal advice, in particular whether private parties' incentives to seek advice when contemplating how to behave or when litigating are socially appropriate. (For example, do parties tend to spend too much or too little to pursue their interests in adjudication?)

Largely separate discussions will be offered for three important legal contexts—the determination of damages, of liability, and of future rights and obligations—because the analysis differs substantially for each. Part II examines a scenario in which different acts cause different levels of harm. While the average level of damage is known, or can be ascertained at low cost, the particular level of damages in a given case can be determined only after more substantial expenditures are made. The question is whether it is worth the cost to measure damages more accurately.

Accurate damage determination may improve individuals' incentives to behave properly.² If I contemplate committing an act that is unusually harmful, I will be more careful or more likely to refrain from the act if I will be held responsible for the true, higher level of harm, rather than for the lower, average level of harm for the class of acts. (Similarly, if my act is less harmful than is usual, holding me liable for an accordingly low amount will prevent excessive deterrence of my activity or avoid creating excessive incentives to be cautious.) In contrast, if at the time I act I am unaware whether my act will cause an atypically high or low level of harm, knowledge that an adjudicator will determine harm precisely *ex post* will not cause me to adjust my behavior accordingly. (If my act creates a risk of injuring a pedestrian and I have no way of knowing how serious a particular injury will be, making the damages I must pay depend precisely on the details of the actual injury cannot affect how carefully I act.) Thus, greater accuracy is valuable only to the extent it involves dimensions about which individuals are informed at the time they act. Relatedly, Part II explores how greater accuracy in adjudication influences individuals' incentives to become better informed before acting. Finally, Part II considers whether parties' incentives to present infor-

² This part also examines briefly how accuracy affects the compensatory function of damages. See Subsection II.A3 *infra*.

mation concerning damages in adjudication are socially excessive or inadequate.

Part III addresses accuracy in determining liability. The context examined is one in which there may be errors: individuals who truly committed acts causing harm—referred to as the “guilty” for convenience³—may mistakenly be exonerated (false negatives), and those who actually did not commit harmful acts—the “innocent”—may mistakenly be sanctioned (false positives). Such errors reduce the law’s deterrent effect. The existence of false negatives means that those who commit harmful acts are less likely to be sanctioned, while false positives increase expected sanctions for innocent behavior. Both effects reduce the disincentive to commit harmful acts.

Thus, one benefit of accuracy is that it is a means of increasing deterrence. The analysis considers what combination of expenditures to increase accuracy and other methods—raising enforcement effort and increasing sanctions—enhances deterrence in the least costly manner. In addition, the analysis demonstrates that greater reliance on accuracy allows a given level of deterrence to be achieved while imposing sanctions less often, which is beneficial whenever sanctions themselves are socially costly (as with imprisonment). The analysis also considers the relationship between accuracy, the level of sanctions, and concerns for sanctioning the innocent. The main result is that, when error is greater, an enforcement policy that relies on higher rather than lower sanctions may be appropriate.⁴ Other topics addressed in Part III are how the burden of proof should be incorporated into the article’s analysis and whether individuals’ incentives to present information in adjudication are socially appropriate.

In Part IV, the analysis shifts to situations in which the purpose of adjudication is to determine either future behavior (such as in licensing procedures) or eligibility for future public benefit payments. Because the analysis of Parts II and III is concerned with incentives for *ex ante* behavior, it is largely inapplicable in this setting. (When future rights and obligations are determined as a consequence of prior behavior, both the analysis of Part IV and that of the prior parts will be relevant.) Determining

³ The analysis applies to the civil context as well.

⁴ The reason, briefly, is that, in achieving a given level of deterrence, higher sanctions permit enforcement effort to be reduced, with the result that sanctions are mistakenly imposed less often; moreover, when enforcement effort is reduced, raising accuracy (further reducing mistakes) becomes cheaper because the required additional expenditures need be made in fewer cases.

the value of accuracy with regard to future events consists largely of identifying the social benefits associated with different outcomes. For example, with entitlements to social security disability payments, one would be concerned with the benefit of supporting the truly disabled, the implicit cost of public expenditures, and the extent to which giving payments to those not truly disabled produces benefits that fall short of the cost of public funds. While this formulation does not appear to deviate substantially from conventional understandings, it is suggested that arguments about accuracy in this context often confuse changes in accuracy with implicit changes in the burden of proof. This part also indicates how incentives for individuals to pursue their claims tend to be socially excessive.

Part V examines concerns that fall outside a conventional economic analysis, particularly those related to the fairness of adjudication. The discussion suggests that some such concerns, notably the problem of mistakenly sanctioning the innocent (whether imprisoning the innocent or requiring individuals not truly liable to make payments to plaintiffs), are already reflected to a substantial extent in the economic analysis. It is explained that, if one's concern about mistakenly imposed sanctions were greater than the economic analysis credits, the implications for legal rules are not always what they appear to be. In addition, Part V addresses the extent to which process values have significance independent of effects on results (as through improving accuracy by providing more information to decision-makers). The discussion in this part is incomplete, in part by design, as the focus of this article is on economic effects, and in part by necessity, as frequently invoked concerns about fairness are rarely sufficiently specified to allow them to be applied and assessed.

II. ACCURACY IN THE ASSESSMENT OF DAMAGES

This part considers efforts to increase accuracy in specifying damages to be awarded in adjudication. The analysis in Section *A* emphasizes the relationship between the degree of accuracy that is anticipated in ex post adjudication and how individuals choose to behave ex ante, in light of the sanctions they expect to be applied to their conduct. It takes into account how the accuracy of adjudication influences individuals' incentives to become better informed before they act. In Section *B*, the analysis is applied to various legal settings. Section *C* extends the discussion to address parties' incentives to present information in adjudication (thereby increasing accuracy).

A. *Accuracy and Ex Ante Information*⁵

1. How the Value of Accuracy Depends on Individuals' Information at the Time They Act

Consider a scenario in which individuals' contemplate committing acts deemed to be harmful. (Such acts may include, for example, torts that involve only a risk of causing harm, breaches of contracts, or violations of intellectual property rights of others.) Acts in a given class are known to cause a particular level of harm, on average; some acts cause more harm than average and others less. (For example, victims' injuries may differ in severity, and given injuries may impose different costs, depending on characteristics of the victim.)⁶

To simplify the discussion, assume that when they cause harm injurers are always liable for damages and the only question is the extent of damages.⁷ The adjudicator⁸ has two choices: award damages equal to the

⁵ Most of the analysis in this section is demonstrated in a formal model in Louis Kaplow & Steven Shavell, *Accuracy in the Assessment of Damages* (Working Paper No. 4287, National Bureau of Economic Research 1993) (hereinafter *Accuracy and Damages*). Related issues are analyzed formally in Louis Kaplow, *The Optimal Complexity of Legal Rules* (Discussion Paper No. 97, Harvard Law School, Program in Law and Economics 1991). Some of the points in this section have been touched on in investigations of legal advice in litigation. See Stephen McG. Bundy & Einer Richard Elhauge, *Do Lawyers Improve the Adversary System? A General Theory of Litigation Advice and Its Regulation*, 79 *Calif. L. Rev.* 313, 381–82 (1991); Louis Kaplow & Steven Shavell, *Legal Advice about Acts Already Committed*, 10 *Int'l Rev. L. & Econ.* 149, 158–59 (1990); Louis Kaplow & Steven Shavell, *Legal Advice about Information to Present in Litigation: Its Effects and Social Desirability*, 102 *Harv. L. Rev.* 567, 588–90, 597–98 (1989) (hereinafter *Legal Advice*).

⁶ Considering victim's incentives, which must be ignored in this preliminary investigation, would make optimal rules more complicated (as when one allows a defense of contributory negligence) or imperfect (strict liability, even with a contributory negligence defense, leaves victims with excessive incentives with regard to their activity level). See, for example, Steven Shavell, *Strict Liability versus Negligence*, 9 *J. Legal Stud.* 1 (1980). Once rules are imperfect to begin with, the value of accurately applying them will differ. One suspects the value may fall. In the simple case in which the rule is strict liability with no defense, awards based on actual rather than average harm would reduce victims' incentives to minimize their expected harm; if this effect were more important than any beneficial effect of accuracy on injurers' behavior, greater accuracy would be undesirable. (A decoupling approach, in which injurers paid damages reflecting actual harm but victims received compensation equal to average harm may avoid this problem.)

⁷ Among the issues ruled out is that concerning incentives to sue and frivolous litigation. It may be that accuracy would discourage frivolous suits, although greater accuracy increases litigation costs, which in some contexts might encourage frivolous suits. See, for example, Lucian Bebchuk, *Suing Solely to Extract a Settlement Offer*, 17 *J. Legal Stud.* 437 (1988); Avery Katz, *The Effect of Frivolous Lawsuits on the Settlement of Litigation*, 10 *Int'l Rev. L. & Econ.* 3 (1990); David Rosenberg & Steven Shavell, *A Model in Which Suits Are Brought for Their Nuisance Value*, 5 *Int'l Rev. L. & Econ.* 3 (1985). See also note 184 *infra* (effect of accuracy on incentive to file claims for public benefits).

⁸ For the present analysis, this may be a judge, a jury, an arbitrator, or whoever.

average harm for the class, or make an inquiry into harm in the particular case, in which event damages will depend on the actual level of the harm.⁹ This inquiry entails some cost.¹⁰ (For the moment, it is immaterial whether the inquiry is undertaken directly by the adjudicator, by a government agency, or by the parties.)¹¹

This subsection examines the value of accuracy with regard to influencing actors' behavior for various assumptions about the information individuals have at the time they act. The influence of accuracy on incentives to become informed *ex ante* and the value of accuracy with regard to compensating victims will be considered in later subsections.

a. *When Individuals Cannot Anticipate the Actual Level of Harm.* Assume that individuals, at the time they decide how to act, know only the average level of harm for the type of act they will commit but not the actual harm their act will cause.¹² Then it is apparent that greater preci-

⁹ Thus, it is assumed that "inaccurate" decisions are unbiased, rather than involving a systematic and therefore predictable over- or underestimate of the harm caused by a type of activity. Instances in which less accurate assessments of damages may involve systematic error are considered in Subsections A1c, A4, and B3. (If the problem is simply that, say, damage awards are 10 percent too low in a class of cases, one might cheaply correct the problem by using a damage multiplier rather than by spending more in each case to improve estimates.)

Relatedly, the assumption is that it is easier (see note 10 *infra*) for the adjudicator to learn average harm than the actual harm. In this case, there are three possibilities—actors *ex ante* know only average harm (considered in Subsection a), both average and actual harm (Subsection b), and only actual but not average harm (which is an interpretation of the analysis in Subsection 4). When instead the actual harm is easiest for the adjudicator to observe, it usually will be clear that damages should simply equal actual harm, regardless of whether actors *ex ante* know actual harm or only average harm (the latter being noted in Subsection B3). (A caveat may arise when defendants are risk-averse, as discussed in Subsection A3, or on account of effects on settlement, as indicated in note 97 *infra*.)

¹⁰ The simple story examined here can be understood in a manner that captures a range of more realistic situations. One might imagine that, after an initial inquiry, a tribunal has some information on the actual harm in a particular case but that uncertainty nonetheless remains. At that point, it has an estimate of the actual harm, which can be taken to be the "average" for purposes of the analysis in the text. (Thus, the "average" need not be across some broad class of acts about which the adjudicator may know little but, rather, can be understood as simply the best estimate before undertaking additional investigation.) The tribunal could, however, engage in further inquiry—at some cost—to refine its estimate for the particular case. This refined estimate, even though itself imperfect, can be taken to be the "actual" harm for present purposes. See Subsection c *infra*.

¹¹ The only effect is that the optimal damage award depends on who pays for the inquiry. In this simple case, if the defendant pays for the inquiry, the optimal damage award if an inquiry has been undertaken simply equals the actual harm. If the government undertakes the inquiry, the optimal damage payment must be increased by the cost of the inquiry. See, for example, A. Mitchell Polinsky & Steven Shavell, *Enforcement Costs and the Optimal Magnitude and Probability of Fines*, 35 J. Law. & Econ. 133 (1992).

¹² It sometimes will be the case that individuals would initially know the actual harm they might cause but not the average that an adjudicator might estimate, perhaps because

sion *ex post*, in adjudication, is a waste of resources because information learned later cannot improve the earlier decision.

For concreteness, suppose that the contemplated act, using a toxic substance, may cause harm of 5, 10, or 15, each with equal probability, if a storage tank leaks. The average harm is 10. If liability always equals 10, the decision whether to use the substance and about how careful to be in preventing leaks will reflect that, if a leak occurs, liability will be 10. If, instead, damages would equal 5, 10, or 15, depending on the actual harm that results, the analysis is the same, for it is assumed that individuals do not know the actual harm *ex ante*. Since they will be held liable for damages of 5, 10, or 15 with equal probability, their expected liability is 10, which will induce the same behavior as if the average harm were used as the basis for damages.¹³ (This assumes that individuals care only about expected liability and not its variance—that is, that individuals are risk-neutral. Risk aversion is considered in Subsection 3.)¹⁴

Thus, greater accuracy has no effect on behavior, while it entails a positive resource cost. As a result, accuracy is of no value and greater accuracy is undesirable.

b. *When Individuals Do Anticipate the Actual Level of Harm.* Assume that individuals, at the time they act, do know the actual level of harm their particular acts will cause.¹⁵ If damage payments reflect the level of harm caused, behavior will be improved. For example, if the benefit from using the toxic substance is 8, it will be used if the actual harm and thus the damage award will be 5, but not if the harm and

they are unsure of what evidence the adjudicator initially will have in making its estimate (see notes 9 & 10 *supra*). The scenario examined in Subsection 4 can be interpreted as including this case. See also Subsection B3 *infra* (possibility that adjudicator must inquire into actual harm because it does not know an individual's average harm).

¹³ The argument in the text involves an oversimplification. If expenditures will be made *ex post*, then they affect behavior *ex ante* either because the actor will anticipate having to make the expenditures or because they will be reflected in an optimal damage award. See note 11 *supra*. Thus, behavior would differ. (More acts would be deterred; greater precautions would be undertaken.) But the primary conclusion that accuracy is undesirable is unaffected, for whatever behavior is produced with accuracy could have been produced without accuracy by simply raising the damage award from 10 to 10 plus the cost of accuracy. That would produce precisely the same behavior as with accuracy but would not involve the expenditure of resources.

¹⁴ See also note 32 *infra*.

¹⁵ Accuracy also will tend to be valuable when individuals do not know the actual level of harm but take some action—not observed by the adjudicator—which affects the magnitude of actual harm in a known manner. (Thus, actors *ex ante* may know more about the harm that will arise than the adjudicator.) Then damages equal to average harm will give actors insufficient incentive to reduce expected harm. See Kathryn E. Spier, Settlement Bargaining and the Design of Damage Awards, J. L. Econ. & Org. (in press); Subsection B3 *infra* (providing analysis similar to what would be applicable in such a case).

damage award will be 10 or 15. In contrast, if the damage award will be 10, to reflect the average harm, the substance would not be used even when the actual harm is only 5. Thus, accuracy avoids excessive deterrence. Similarly, if the benefit from using the substance were 12, it would be used even when the actual harm would be 15 if damages were based on the average harm of 10, but not if damages equaled the actual harm of 15. Accuracy thereby provides efficient deterrence that otherwise may be lacking.¹⁶

The question remains whether this improvement in behavior on account of accuracy is sufficiently desirable to justify the cost of greater accuracy. If most leaks caused harm of 10 and accuracy were expensive, the cost of establishing whether there were exceptions in each case would exceed any benefit in improving behavior. Similarly, accuracy would be undesirable if decisions concerning use of the substance would be unaffected in any event. (This would occur either if most users had benefits less than 5, in which case they would not act regardless of whether damages were estimated accurately, or if most users had benefits exceeding 15, in which case they would act in any event.)¹⁷ In contrast, accuracy will be valuable when it is cheap and the effect on use of the substance would involve substantial benefits.

c. *The Optimal Degree of Accuracy.* When individuals are informed about the level of harm that their acts might cause, the optimal level of accuracy will be a matter of degree. It would no doubt be undesirable to treat auto and aircraft collisions as a single group because the average level of harm in each category differs greatly; moreover, the cost of distinguishing these categories is extremely low. Similarly, one would wish to distinguish cases in which cars hit pedestrians from those in which cars crush strands of grass on someone's lawn. It may be very costly, however, to determine whether a victim's loss in future earnings will be 30 percent or 40 percent of his previous potential, while such a difference may have little effect on precautions.

In addition, the more refined the damage inquiry, the less likely it is

¹⁶ The text oversimplifies because behavior with accuracy would also reflect the cost of the accuracy, as discussed in note 13. Accounting for this would not fundamentally change the analysis. (For example, if accuracy costs 1, the expected total costs for individuals who commit acts causing harm of 5, 10, and 15 would be 6, 11, and 16, respectively, and the example in the text would illustrate the same potential effects of accuracy.)

¹⁷ If they would not act, of course, the litigation costs would not be borne in such instances. If one considered instead a scenario in which individuals would act regardless and the only question were how much care they would take, the analogous argument is that, if care would be little influenced by the differences in actual harm, it would be wasteful to be more accurate.

that any difference in outcome will be anticipated at the time individuals decide how to act. One who drives into a pedestrian in a crosswalk may cause harm of thousands or millions of dollars. But the degree of harm will not be known to the driver in advance. Thus, a practice of using averages—say, for types of accidents or types of injuries—may have little effect on behavior even when the range of potential difference in actual harm is vast.¹⁸

2. The Degree of Ex Post Accuracy and Individuals' Incentives to Acquire Information Ex Ante

a. *How Accuracy Ex Post Induces Individuals to Acquire Information Ex Ante.* The analysis in Subsection 1 assumes that individuals, at the time they decide how to act, either are or are not informed about an aspect of the harm they might cause. Often, however, the extent to which individuals are informed will be a matter of choice. For example, one who contemplates using dynamite may not know very precisely how dangerous it is but may be able to consult experts who are more familiar with the extent of the danger from various uses. The more individuals are willing to spend, the more information they can acquire.¹⁹

The central point of this subsection is that the extent to which individuals will choose to become more informed ex ante, when contemplating how to act, will depend on the degree of accuracy they expect in adjudication, ex post. An individual will see no value in making an expenditure to learn whether her act will cause harm of 5, 10, or 15 if she knows that a court will award 10 in any event.²⁰ Only if she anticipates that the court will learn the level of harm more precisely, and make damages reflect the actual harm, will she have an incentive to learn the actual level of harm ex ante.

¹⁸ Relatedly, if the driver knew in advance that substantial harm would be caused, she would have been more careful. And if she knew that reckless driving in her lifetime would in fact produce trivial harm—that is, if she knew in advance that she would be extremely lucky—she would have been less careful. The point is that such differences in outcomes are often unknown in advance, in which case they do not affect behavior.

¹⁹ One can think of the two cases in Subsection 1 as representing extremes with regard to ex ante information costs. The case in which individuals were assumed to be informed is similar to one in which information is extremely cheap, while the one in which they are not informed might be one in which information is prohibitively expensive. (In many instances, of course, predicting future events with high precision may be virtually impossible; one can think of the information cost as infinite, in which case individuals would never acquire the information in any event.)

²⁰ If the actor and potential victim are in a contractual relationship, there may be an incentive with regard to negotiation of price and adjusting behavior. See note 41 *infra* (noting the Hadley v. Baxendale problem).

Whether individuals ultimately acquire the information is another matter. Individuals will acquire information when the benefits of being informed exceed the costs of the information. The benefit of being informed is that one can better adjust one's behavior in light of actual legal consequences. For example, if there is a probability that one will learn that actual harm, and thus expected liability, is very high, and one would choose not to act in that instance, then the benefit of information would be the difference between the net value of not acting and that of acting in such an instance, weighted by the probability one expects to learn that actual harm is very high.²¹ Similarly, one might learn that harm is lower than anticipated and in that instance commit an act that one otherwise would have refrained from doing.

To illustrate, suppose that the use of a toxic substance will cause harm of 5, 10, or 15, each with equal probability, and that the user does not know the actual harm that would be caused. If the benefit from using the substance is 12, it would be used in the absence of information because the expected liability is 10. If one acquired information on actual harm, the substance would not be used in the event that harm was 15.²² This produces a gain of 3—liability of 15 is avoided while the benefit of 12 is forgone. Because the probability of this outcome is 1/3, the expected value of information is 1.²³ Thus, information would be acquired if and only if the cost of the information were less than 1.

Information *ex ante* is valuable only if what is learned will be reflected in awards *ex post*. Similarly, as noted in Subsection 1, accuracy *ex post* is valuable only if individuals also have accurate information *ex ante*; that statement can now be interpreted to include both the case in which individuals already have the necessary information and that in which they will be induced to discover it. And individuals will be induced to learn information *ex ante* only if their benefits from adjusting behavior exceed the cost of the information.

²¹ The value of not acting can be taken to equal zero as a reference point. The value of acting would be the benefit of the act minus the expected sanction; in this instance, the value is negative. (If the difference were positive, it would be desirable to act even if one learned that harm was very high. Then the individual would have had no reason to acquire the information.)

²² When the user learns that actual harm is 5 or 10, the substance will be used, for the benefit is 12, but this is the same result as when information is not acquired. Thus, acquiring information affects behavior only in the case in which actual harm is 15.

²³ An analogous example can be constructed for the case in which one is induced to act by favorable information. If the benefit from using the substance were 8, one would not use it without first obtaining information about actual harm, because the expected liability is 10. If the information indicates that actual harm is only 5, the substance would be used, for a gain of 3 and an expected gain of 1.

b. *The Social Value of Accuracy Ex Post.* Having established when accuracy is valuable, it remains necessary to consider whether the value is worth its cost. In Subsection 1b, when individuals simply were assumed to be informed, accuracy ex post was desirable when the benefit of the improvement in behavior was greater than the cost of more accurate adjudication ex post. In the present context, the benefit with regard to behavior must exceed the sum of the cost of accuracy ex post and the cost of individuals' becoming informed ex ante.²⁴

c. *Whether Individuals' Incentives to Acquire Information Ex Ante Are Excessive or Inadequate.*²⁵ It is noted in Subsection a that individuals acquire information when the anticipated benefits to themselves exceed its cost. This subsection explores whether, ex ante, individuals will tend to acquire more or less information than is socially appropriate (taking as given that adjudication will involve the expenditure of resources to reach accurate results ex post).

First, consider a benchmark case. Assume that liability is strict, the only uncertainty concerns damages, and individuals properly estimate the value of information to themselves.²⁶ Then, if the information concerns how a court will properly assess harm, individuals' incentive to acquire information will be socially appropriate.

The reason incentives are proper in this simple context relates to familiar properties of a rule of strict liability.²⁷ When individuals acquire information about the true level of harm, and expected damages equal expected true harm, they will be induced to behave appropriately. Relatedly, the private benefit from their change in behavior will equal the social benefit. (The private benefit will consist of gains or losses from modifying one's activity, which are assumed to be social gains or losses as well. Changes in expected liability payments will equal changes in expected harm, which is a correct measure of the external social cost of one's activity.) Finally, individuals' private cost of information involves real expenditures of resources (usually, time or expenditures to purchase experts' services), which will thus be a proper measure of the social cost of information. Therefore, when individuals compare their private

²⁴ In contexts in which harm is probabilistic—as is typical with accidents—note that ex ante costs will be borne more frequently than ex post costs.

²⁵ This topic is the subject of Louis Kaplow & Steven Shavell, *Private versus Socially Optimal Provision of Ex Ante Legal Advice*, 8 J. L. Econ. & Org. 306 (1992).

²⁶ More precisely, the assumption is that their probability distribution of what they might learn is correct—rather than, say, biased toward expecting that what they learn will be systematically more positive than is actually the case.

²⁷ See, for example, Shavell, *supra* note 6.

benefits and costs, their calculus will precisely reflect social benefits and costs, and their incentive to become informed will be correct.

Before discussing contexts in which this conclusion may not hold, it is useful to explore the implications of social and private incentives to acquire information being properly aligned. On one hand, alignment implies that there is no case for taxing or subsidizing individuals' purchases of such information²⁸—which often may consist of legal services. On the other hand, alignment of *ex ante* incentives leaves open the question whether the degree of accuracy *ex post* is appropriate. The analysis of private incentives to acquire *ex ante* information *takes as given* the assumption that damage awards will be accurate with regard to the type of information they might acquire. It remains the case that such accuracy will be desirable only when the benefits from improved behavior exceed the *ex ante* costs of information acquisition plus the *ex post* costs of greater accuracy. That private incentives are correct establishes only that the benefits from improved behavior exceed the *ex ante* information costs. Whether the difference between these benefits and costs (summed across the relevant population of individuals)—which is the net *ex ante* benefit—exceeds the *ex post* cost of greater accuracy is another matter.

Finally, it should be noted that private incentives to acquire information *ex ante*, even given the level of accuracy in adjudication, need not be socially appropriate. In the context explored in Subsection 4 below, in which *ex ante* information concerns errors adjudicators will make *ex post*, private incentives tend to be excessive. Similarly, when a negligence-type rule is employed, incentives may be excessive.²⁹

For present purposes, the possibility that private incentives to acquire information may be socially excessive, or inadequate, complicates a practical inquiry into the value of accuracy but not the fundamental elements. It is still the case that one must weigh the benefits with regard to behavior against the costs of *ex ante* and *ex post* information. The only difference is that assessing the benefits with regard to behavior and the costs of *ex*

²⁸ This is not to deny that there may be a role for government intervention. If there are public good aspects to such information, it may be efficient for the government to collect and publish it—for example, by producing or subsidizing the private creation of databases with accident statistics. Aside from such acts, however, individuals' incentives to learn about what is in such databases and how the information may be applied to their behavior would be appropriate. See also Subsection B2c *infra*.

²⁹ See Kaplow & Shavell, *supra* note 25. A “negligence-type” rule is meant to include any rule in which one is not liable for any damages if one's behavior falls within a protected zone, but one is liable if one's behavior is outside the protected zone. This would include, for example, the negligence rule in tort law, contract rules finding breach only when behavior crosses some threshold, and public regulations that impose penalties only for violating standards that permit some harm-producing activity.

ante information may be more complicated when ex ante incentives to acquire information are inappropriate.

3. Compensatory Objectives and Risk Aversion

When parties are risk-averse—that is, when uncertain losses are more detrimental than a certain loss with the same expected value—the level of actual damage awards will affect their welfare directly, in addition to affecting their behavior.³⁰ To illustrate this, recall the example in which actual harm may be 5, 10, or 15, each with equal probability.

Consider first the effect with regard to injurers (defendants), who have been the focus of the analysis thus far. In the scenario examined here, inaccuracy corresponds to making defendants pay the average harm caused by their type of act (10), while accuracy involves each defendant paying the actual harm caused by her particular act (5, 10, or 15). By construction, the expected payment (the mean) is the same in both instances (10). But the variance differs: there is no variance if the average harm is paid by all injurers, while there is positive variance if each injurer pays the actual harm she causes (the standard deviation is approximately 4.1).³¹ As a result, greater accuracy in adjudication increases the risk to which defendants are exposed. On this account, accuracy is less desirable than otherwise.³² Thus, when individuals are uninformed ex ante, so that there is no behavioral benefit, accuracy would be undesirable even if the greater accuracy were free!³³

For plaintiffs, the effect would tend to be the opposite. Assume that plaintiffs' damages compensate for pecuniary losses (or nonpecuniary losses, such as physical injuries, that are fully restored by expenditures of money). Then, if they are risk-averse, the optimal amount of compen-

³⁰ The possibility that one may care about accurate compensation independent of concerns for imposing risk is considered in Subsection VA1.

³¹ More generally, there is positive variance and thus risk in both instances if an injurer's act causes harm only with a probability, for then when no harm arises, the payment is zero. Nonetheless, the variance is greater when, in instances where harm occurs, individuals pay for actual harm rather than the average harm for such instances.

³² A complete analysis would be more complicated because, when defendants are risk-averse, optimal damages would no longer equal actual harm. Rather, they would involve some compromise between optimally controlling behavior ex ante and minimizing the imposition of risk. (Relatedly, the effect of damages on behavior differs when individuals are risk-averse.) The statements in the text would, nonetheless, remain valid even if such complications were taken into account.

³³ If individuals become informed before they act, the risk of a more accurate system is borne at the stage at which they become informed.

sation equals the actual loss.³⁴ Hence, if damages equal average harm for classes of plaintiffs rather than actual harm, some plaintiffs (those with harm of 15) will be undercompensated and others (with actual harm of 5) will be overcompensated (as both types receive damages of 10). Therefore, plaintiffs will bear risk when results are inaccurate, but not when they are accurate.

Thus, taking risk aversion into account could make accuracy more or less desirable than otherwise, depending on whether plaintiffs' or defendants' risk aversion were more significant. Of course, all these conclusions are mitigated to the extent parties are insured.³⁵

4. Variation: Reduction of Error That Is Costly to Predict³⁶

The analysis in this section has considered the situation in which an inaccurate adjudication is one that involves awarding damages equal to the average harm for a class of activity, while accurate adjudication involves damages equal to actual harm. Consider now a different type of inaccuracy. Suppose that, if the adjudicator is inaccurate with regard to a particular act, it awards too high a level of damages in some instances and too low a level of damages in others but that, on average, its awards are correct. To be concrete, consider the case in which true harm is 100,

³⁴ If injuries are nonpecuniary, optimal compensation need not equal actual loss. For example, if individuals incur substantial pain and suffering, but this does not alter their marginal utility of wealth, the optimal compensatory payment on account of risk aversion is zero. See, for example, Steven Shavell, *Economic Analysis of Accident Law* 228–35, 245–54 (1987). As between payments that will equal some class average and payments that will reflect one's actual pain and suffering, risk-averse individuals would prefer the former. In this instance, therefore, both risk-averse plaintiffs and risk-averse defendants would prefer the less accurate system, with damages equal to average harm rather than actual harm.

³⁵ To the extent insurance coverage is nearly complete, the behavioral benefits of liability rules will tend to be less, making the very presence of a system of adjudication that makes injurers pay damages less likely to be desirable; relatedly, an expensive and precise system will be less attractive than a cheaper, imprecise one. Accuracy may still be valuable to the extent that insurance companies can identify injurers in advance as those likely to cause more or less harm than average. (That is, the question becomes whether insurers, rather than whether actors, know actual harm *ex ante*.) For then premiums will be adjusted for such individuals, which may affect behavior. (For example, those whose acts cause more injury and thus who must pay high insurance premiums may forgo the activity. Or businesses will reflect these costs in prices, leading to a more efficient allocation of resources.) See also Bharat Sarath, *Uncertain Litigation and Liability Insurance*, 22 *RAND J. Econ.* 218 (1991) (suggesting that liability insurance may be undesirable if there is error in determining liability).

³⁶ This scenario is modeled in Kaplow & Shavell, *supra* note 25, at 312–16, although that article does not address the issue of when greater accuracy is desirable.

but 10 percent of the time damages are mistakenly determined to be 75 and 10 percent of the time they are 125. Expenditures making the adjudication accurate eliminate such errors.³⁷

When an individual cannot anticipate the error a court will make in her particular case, accuracy is of no value. The reason is that, in such instances, she will act based on the expected award (100), which is assumed in this instance to equal the actual harm. (The difference with the prior analysis is that here behavior is ideal when individuals are uninformed about the precise outcome of the adjudication.)

When an individual does anticipate the error a court will make, she may behave incorrectly. For example, if she anticipates that the court will err on the side of awarding damages that are too low (75), she will take too little precaution or may be induced to undertake an activity the benefits of which exceed the expected liability payment but are less than the true expected harm (for example, an activity with a benefit of 90). If the court were to make the expenditure necessary to eliminate the error, however, then the expected liability payment would equal the true expected harm (100), so behavior would be appropriate. Thus, as before, accuracy is valuable in the case in which individuals have information *ex ante*. (The difference is that here individuals' information concerns court error in the case in which adjudication is inaccurate, rather than true harm in the case in which adjudication is accurate.)

Finally, suppose that individuals do not know the error an inaccurate court will make but can predict it if they make an expenditure on information.³⁸ In this case, in contrast to that in Subsection 2c, such information is undesirable (and thus, *a fortiori*, less socially valuable than its positive cost). After all, uninformed individuals behave properly (because their expected liability equals actual harm), while informed individuals may behave improperly (when they learn that actual liability differs from actual harm).³⁹ Thus, in principle, one might wish to discourage the acquisition of such information. Alternatively, this problem would be a reason to increase accuracy: when adjudicators are accurate, there is no longer any error to learn. Individuals are therefore induced to behave better, and they no longer have an incentive to make wasteful *ex ante* investments to

³⁷ The analysis would be qualitatively the same if the expenditures reduced either the probability or magnitude of such errors rather than eliminating them entirely.

³⁸ For example, they may hire a lawyer who can predict typical jury biases or who knows that rules concerning consequential damages omit certain components of harm.

³⁹ Another possible dimension of improper behavior may involve *ex ante* or *ex post* activity designed to mislead the adjudicator. Greater accuracy may discourage such costly efforts, or it might lead individuals to substitute more costly techniques that will remain undetected even by more accurate adjudication.

predict errors (that no longer will be made). Thus, greater expenditures on accuracy *ex post* are justified when the cost is exceeded by the benefit from improved behavior plus the reduction in *ex ante* information acquisition costs.⁴⁰

In summary, as before, accuracy tends to be valuable to the extent individuals are informed *ex ante*. The primary difference concerns the incentive to become informed. In the original scenario addressed in Subsections 1–3, more accuracy *ex post* induces greater information acquisition *ex ante*; here, more accuracy reduces the incentive to acquire information *ex ante*. In both instances, greater accuracy tends to improve behavior if individuals are informed *ex ante* but has no effect on behavior if they are uninformed (and will not be induced to become informed) *ex ante*.

B. Applications

This section applies the analysis in Section A to various legal settings. Some of the applications are generic, consisting of comments on the type of setting or legal question to which the analysis applies. Others are specific and are offered because of their independent significance as well as their value in illustrating the analysis more concretely.

1. Generality with Regard to Areas of Law and Types of Rules

The scenarios outlined and analyzed in Section A are most directly interpreted as involving accident law and a rule of strict liability. But the analysis is more general in a number of dimensions. First, much of law, particularly public regulation, is like tort law in that it attaches sanctions to activities that may cause harm to third parties. This includes most substantive rules outside of the contractual context.

Second, in contractual settings, the scenario is directly applicable to questions of breach. Whether rights are determined by background law or terms supplied by the parties, the question remains how accurately an adjudicator should determine harm.⁴¹ (The concrete question might

⁴⁰ Note that in this instance a more accurate legal system may reduce total legal costs. This would be the case if the reduction in *ex ante* expenditures on legal advice itself exceeded the greater legal costs from more accurate *ex post* adjudication.

⁴¹ One particular application involves the problem of foreseeability of unusually high damages, as in the famous case of *Hadley v. Baxendale*, 9 Ex. 341, 156 Eng. Rep. 145 (1854). It is familiar that, if the promisor is not informed about when the harm from breach is severe, there is no benefit with regard to the promisor's behavior in making damages high in such cases. See, for example, Richard A. Posner, *Economic Analysis of Law* § 4.9 (4th ed. 1992); Lucian Arye Bebchuk & Steven Shavell, *Information and the Scope of*

involve, for example, the measurement of expectation damages.) Also note that parties often provide for alternative dispute resolution in their contracts. This merely transfers the question of the appropriate degree of accuracy to another forum. Parties might specify the level of accuracy with which they would like an arbitrator to resolve disputes or they may choose a particular provider of dispute resolution services that is known to provide a desired level of accuracy (at a cost that reflects how accurate an outcome is usually produced).⁴² Or, they may specify liquidated damages, making an inquiry into actual harm unnecessary.

Third, the scenario was presented in a manner suggesting that adjudication concerned the level of damages rather than, say, whether an activity is of a type or was conducted in a manner that subjects one to liability. But the scenario can readily be interpreted differently. Suppose, for example, that acts in a given class are of two actual types, one harmless and another causing a certain degree of harm (denoted H). An accurate adjudication would apply a rule of no liability to the harmless type of act and liability (with damages of H) to the harmful type of act. An inaccurate adjudication—one that failed to distinguish the two types of acts—might award damages of less than H to all acts (where the amount of damages equaled the average harm— H times the proportion of acts in the class that are of the harmful type). But this is functionally equivalent to saying that there is liability for all acts and, if the court is accurate, damages equal true harm, which is zero for the harmless type of act. As a result, interpreting the scenario as involving uncertainty over two possible levels of harm (rather than a potentially unlimited range of possibilities), when one possible level is zero, indicates that the analysis is applicable to accuracy in categorizing activity.

An important qualification is that the analysis assumed in all contexts that the legal rule took the form of strict liability—that is, the injurer pays the victim's harm even if the injurer behaved reasonably. Under a negligence-type rule, the analysis may be more complicated, as small differences in behavior might be responsible for great differences in liability payments.⁴³ The results, nonetheless, are qualitatively similar. If indi-

Liability for Breach of Contract: The Rule of *Hadley v. Baxendale*, 7 J. L. Econ. & Org. 284 (1991).

⁴² Those designing alternative dispute resolution mechanisms, such as for the American Arbitration Association or in state statutes providing, say, mandatory arbitration for complaints that a new car is a lemon, confront the same questions facing architects of the conventional legal system (with the possible difference that constitutional constraints may be absent).

⁴³ See note 29 for elaboration on what is meant by negligence-type rules. The effects of a negligence rule on behavior in the presence of uncertainty is explored in Richard Craswell

viduals are uninformed *ex ante*, there is no value in an adjudicator having more precise information *ex post*.⁴⁴ If individuals are informed *ex ante*, the anticipation of greater accuracy *ex post* will improve behavior.⁴⁵ Finally, individuals' incentive to acquire information *ex ante* will be affected by the degree of accuracy that will be employed *ex post*.⁴⁶

2. Formulating Legal Commands: Rules versus Standards and Complexity⁴⁷

Two recurring questions concerning how legal commands should be formulated involve the appropriate degree of complexity or detail (should acts and outcomes be distinguished finely or grouped more broadly) and whether commands should take the form of rules or standards (should the content be determined and announced in advance, in a rule, or left to an adjudicator, in a standard). As I have argued elsewhere, it is useful to distinguish these two aspects of formulation and consider them separately. The reason is that two dimensions are involved that often can be varied independently. To illustrate, one could have a simple rule (a penalty of \$10 for any parking violation), a complex rule (a detailed schedule

& John E. Calfee, Deterrence and Uncertain Legal Standards, 2 J. L. Econ. & Org. 279 (1986); Jason S. Johnston, Bayesian Fact-finding and Efficiency: Toward an Economic Theory of Liability under Uncertainty, 61 S. Cal. L. Rev. 137 (1987); Shavell, *supra* note 34, at 79–83, 93–99. As Kahan notes, however, the doctrine of causation may eliminate the characteristic of a negligence rule that results in substantial differences in liability for small differences in behavior, in which case the analysis is closer to that applicable to strict liability. See Marcel Kahan, Causation and Incentives to Take Care under the Negligence Rule, 28 J. Legal Stud. 427 (1989). The effect of these differences with regard to *ex ante* incentives to acquire information is examined in Kaplow & Shavell, *supra* note 25, at 311–12, 316.

⁴⁴ While this argument is more complex because of the manner in which a negligence rule affects behavior in the presence of uncertainty, it still holds. See Kaplow & Shavell, Accuracy and Damages, *supra* note 5, at 18 n.21.

⁴⁵ With a negligence rule, there are scenarios in which straightforward use of greater accuracy could make behavior worse. See Craswell & Calfee, *supra* note 43, at 287 (showing that reducing the variance in estimates under a negligence rule can result in worse behavior). But it will generally be true that if one used the more accurate information optimally (including the possibility of adjusting the rule, as by changing the criterion for due care), one could improve the control of behavior. (It is trivial to ensure that matters do not get worse; if they did, the information could be made legally irrelevant. Usually, however, there would be a way to use the better information to improve behavior.)

⁴⁶ Distortions in the incentive to acquire information *ex ante* under a negligence rule are analyzed in Kaplow & Shavell, *supra* note 25, at 309–12, 314–16.

⁴⁷ The analysis in this subsection largely derives from Louis Kaplow, Rules versus Standards: An Economic Analysis, 1992 Duke L. J. 557. A formal analysis of complexity appears in Kaplow, *supra* note 5. Some of the points concerning rules versus standards appear in Anthony Ogus, Information, Error Costs and Regulation, 12 Int'l Rev. L. & Econ. 411 (1992).

of parking penalties that depend on location, time of day, day of week, and driving conditions), a simple standard (the adjudicator sets the penalty, taking into account only that there was a parking violation), or a complex standard (the adjudicator considers location, time of day, and so on in setting the penalty). This subsection briefly considers the relationship between these aspects of formulation and the problem of accuracy.

a. *Complexity.* The question of the appropriate degree of accuracy in adjudication and that of the appropriate degree of complexity or detail in legal commands are closely related. Consider the scenario introduced in Subsection A1. The question posed was whether it is worthwhile under a rule of strict liability for an adjudicator to inquire into the precise level of harm rather than holding defendants liable for the average harm for the relevant class of acts. Compare this question to one of how detailed a rule for harmful behavior should be. A simple rule specifies that acts in the described class of behavior shall give rise to liability of a stated amount and sets this amount equal to the average harm of acts in the class. A complex rule specifies a damage award for each type of actual case that may arise, with the award equal to actual harm. Under the simple rule, evidence concerning the actual harmfulness of an act would be irrelevant. Under the complex rule, one could not determine what damage amount was applicable unless one first assessed the particular act that occurred. In essence, the simple rule is a command that the adjudicator not invest in more accurate information, and the complex rule is a command that the adjudicator invest in more accurate information. Thus, the specificity of substantive rules is one of the ways in which a legal system indicates how accurate its adjudication is meant to be.⁴⁸

The difference between simple and complex standards involves much the same choice. A simple standard, recall, is one that does not involve the adjudicator making a precise *ex post* inquiry but, rather, making its decision as to the appropriate level of damages in a more general fashion. In contrast, a complex standard is one in which the adjudicator takes into account more factors—which requires more precise information—in determining what damages must be paid. (The difference is that, when standards are used rather than rules, the determination of what penalty corresponds to what set of facts is made by the adjudicator, *ex post*, rather than the promulgator of the rule, *ex ante*.)

Furthermore, when adjudicators operate under standards, as they often

⁴⁸ Of course, a substantive law could be even more detailed: it could specify what evidence counts as proof of an element or, similarly, indicate what damages shall be awarded for various combinations of evidence that may be presented. See note 49 *infra*.

do at least implicitly, they must decide what level of specificity is appropriate. If a law merely commands that those who park improperly shall pay accordingly, the adjudicator, in interpreting the law, must determine how many dimensions in how much detail are to be considered; the more complex the inquiry, the more refined the necessary evidence will be. Thus, a decision on how complex the standard should be is a decision about how accurately the adjudicator should assess behavior.⁴⁹

b. *Rules versus Standards.* Compare a complex—that is, a detailed, precise, accurate—standard to a complex rule with the same content (one that provides for the same level of damages for each particular act that may arise).⁵⁰ How, if at all, would their effects on behavior differ? If individuals were, in either case, unaware of the detail of the law, they would act in the same manner under both formulations, in a manner that reflected the expected liability for a class of acts rather than the actual harmfulness of their particular act. Similarly, if they were aware of the precise consequences that would follow from their activity in either case, they would act the same way, in a manner that reflected the particular harm of their act.

But it is plausible to suppose that individuals' information will differ under the two formulations. In particular, individuals may find it easier (cheaper) to become informed under rules. Because the specification of particular consequences is stated in advance, legal consequences are less costly to determine than under a standard, where by definition the statement of damages for various acts is left to an adjudicator, *ex post*. Thus, even when the content does not differ, individuals may often be more informed about the particulars in rules than in standards at the time they act and thereby conform their behavior more closely to the law's commands.⁵¹

Moreover, it may sometimes be true that even if one compared a less

⁴⁹ One might say that, once an element is deemed to be relevant in a particular manner, there remains the question of how accurately to ascertain that element. But the distinction is largely semantic. In principle, one can state a functional relationship directly from evidence to outcomes: findings on elements are a function of evidence, and outcomes are a function of findings on elements; thus outcomes are a composite of these two functions and can be determined from the evidence. The deemed relationship between evidence and findings can be viewed as part of the system of legal commands just as is the relationship between findings and outcomes. When evidentiary presumptions and proof burdens are stated formally, this is obvious, but the logical relationship remains the same even if the relationships are implicit.

⁵⁰ The validity of this sort of comparison is discussed in Kaplow, *supra* note 47, at 586–96.

⁵¹ The suggestion that some standards may be more accessible to individuals than rules is criticized in *id.* at 596–99.

complex rule to a more complex standard (perhaps one that could be more detailed because hindsight is better than foresight), the former, less complex command may result in behavior that is more precisely in accord with underlying legal norms. The reason is that the standard's added detail may be unknown, and too difficult to predict, *ex ante* (for the very reason that hindsight has yet to be obtained). Thus, the conventional view that there is an advantage in allowing more room for adjudicators to examine context-specific factors, rather than specifying legal consequences in advance under a simpler scheme, exhibits an error in logic. This view implicitly assumes that the more precise *ex post* result will be reflected in *ex ante* behavior; but this will not occur if the ultimate content of the legal command cannot cheaply be predicted in advance (which is often a major premise of those presenting this justification for standards). Moreover, failing to announce (formulate into rules) even simpler aspects of an appropriate outcome may make it more difficult for individuals to take them into account in their behavior, leading to worse results. This analysis is simply an application of the main principle of Subsection A1: accuracy *ex post*, in adjudication, is valuable with regard to improving behavior only to the extent individuals have the relevant knowledge *ex ante*, when they decide how to act.

There is, however, an exception to this maxim. Recall from Subsection A4 that, if errors are made in formulating the law and individuals can predict errors in the outcome of adjudication, their behavior will be worse than if they cannot. Assume that rules, when promulgated, and standards, when applied, contain errors. Errors under rules are easier for individuals to learn and thus more likely to affect behavior in an undesirable manner. In contrast, precisely because individuals are more poorly guided by standards, they are less likely to be influenced (adversely) by errors in the application of standards.⁵²

c. *Predictability of Legal Outcomes.* The arguments in Subsection b are based on the preceding analysis indicating that predictability is desirable when greater accuracy is involved and undesirable when error is involved. These ideas have implications aside from questions literally involving the law's formulation. If one wants the law to be a more precise guide to behavior, individuals need the relevant knowledge. One way to disseminate it is through rules, which state the consequences that attach to various behavior. Another way to disseminate information to guide

⁵² Even this exception has a caveat: if errors under standards are likely, individuals have an incentive *ex ante* to expend resources to predict them. To the extent individuals undertake such expenditures, behavior will be little better under standards, while more (wasteful) expenditures on *ex ante* advice will be made.

behavior is through the outcomes of adjudication. Obviously, one benefit of more accurate outcomes is that individuals in the future will have better guidance concerning their behavior.⁵³

But adjudication tends to be a costly and ineffective way to collect and disseminate information. Often more will be spent adjudicating a single case than would be required to fund a substantial empirical study of a mass of cases. Moreover, if a case settles, typically few individuals learn of the information produced through the litigation. Even if there is a verdict, it is extremely difficult to interpret a black-box pronouncement by a jury, without statements of reasons, weight given to factors, or basic evidentiary findings.⁵⁴

Thus, while adjudication does, at substantial cost, provide some useful information, governments should consider alternatives that involve directly gathering and disseminating information.⁵⁵ For example, if one is concerned about the waste-disposal behavior of dry-cleaning establishments or auto repair shops, study of appropriate methods of disposal and costs of inappropriate methods, followed by distribution of the results, may be an important accompaniment to substantive law regulating the activity. (Of course, such investigation would simultaneously provide the information needed to give appropriate content to the legal commands.) Even if one chose to enforce compliance using a conventional tort regime, there would exist the advantage that, when adjudicators set damages equal to actual harm, individuals would be more likely to know actual harm at the time they decide how to act. (In addition, adjudicators may find that they could assess harm more accurately and more cheaply if such an investigation had been undertaken.)

In contrast, the legal system may function more effectively if its errors and idiosyncrasies are less predictable. Randomly assigning judges and making it difficult to learn how juries actually reach decisions interfere with prediction.⁵⁶ Limiting appellate review of fact-finding, usually justi-

⁵³ This subsection considers only the effect of predictability on ex ante behavior, not on settlement.

⁵⁴ See James F. Blumstein, Randall R. Bovbjerg, & Frank A. Sloan, *Beyond Tort Reform: Developing Better Tools for Assessing Damages for Personal Injury*, 8 *Yale J. Reg.* 171 (1991) (exploring the problem and advocating solutions); Randall R. Bovbjerg, Frank A. Sloan, & James F. Blumstein, *Valuing Life and Limb in Tort: Scheduling "Pain and Suffering,"* 83 *Nw. U. L. Rev.* 908 (1989).

⁵⁵ See note 28 *supra*.

⁵⁶ Prediction of jury behavior is impeded by using general verdicts rather than detailed special questions and by limiting lawyers' abilities to interview jurors afterward. See, for example, *Tanner v. United States*, 483 U.S. 107 (1987) (Federal Rule of Evidence 606(b), which prohibits juror testimony about deliberations to impeach a verdict, applied to testimony about jurors' alcohol and drug use during trial). In addition, jury outcomes are not

fied on unrelated grounds, also has the effect of limiting the knowledge base of legal experts and individuals who contemplate activity in the future. Of course, all such limits also disguise what may be important information reflecting the (at least roughly) accurate findings of adjudication.⁵⁷

3. Inquiries about Ex Ante Knowledge versus Ex Post Consequences

There are two important differences between inquiring into an actor's ex ante knowledge and inquiring into what is known ex post, including the actual consequences of her acts. First, different inquiries may be appropriate as a matter of principle. For example, ex ante knowledge—what individuals actually knew or should have known—may be most relevant for incentives⁵⁸ while ex post consequences may be most relevant for determining appropriate compensation of risk-averse plaintiffs.⁵⁹

Second, different inquiries may involve different costs. In some cases, ex post consequences will be obvious, but ex ante knowledge (actual or what an actor should have known) may be extremely difficult to determine. In others, ex ante knowledge may be straightforward while there is substantial dispute concerning what actually happened. (When an individual commits an act that is readily classified and has a known distribution of consequences, ex ante knowledge may be apparent; yet whether a particular harm was caused by an act or how extensive the harm was to a particular victim may be quite difficult to ascertain.)

Such considerations have an important bearing on accuracy in adjudication. For example, it is often the case that individuals can know at most the probability distribution of consequences of their acts. (The effect of a form of air pollution may be reasonably well known statistically, although it is impossible to predict whether particular molecules that are emitted will ever enter anyone's lungs and what damage they will cause that otherwise would not have occurred.) At the same time, greater accu-

officially reported, making it difficult to match outcomes to case characteristics. Of course, commercial services and grapevines communicate such information, although rather imperfectly. (Each decision is noisy and involves many factors; generalizing across a broad sample of such decisions will thus be difficult, particularly with informal networks in which communications are incomplete and perhaps purposely biased to reflect positively on the lawyer telling the story.)

⁵⁷ Intentionally designing the legal system to make it more difficult for individuals to understand how it operates raises questions of legitimacy and accountability that go beyond the scope of this article, although some such issues are noted in Section VC.

⁵⁸ See Steven Shavell, *Liability and the Incentive to Obtain Information about Risk*, 21 J. Legal Stud. 259 (1992).

⁵⁹ See Subsection A3 *supra*.

racy *ex post* may be extremely expensive. (Determining whether a plaintiff was actually exposed to a defendant's emissions and, if so, the extent to which his illness can be attributed to such exposure may be very costly and unavoidably uncertain.)

In such instances, not only would accuracy be costly, but it would be of no value. Evidence of the harm the defendant caused the plaintiff would be only a fragment of the data concerning the defendant's expected harm. Of course, if the defendant will be held to account accurately for all harm caused, the defendant's expected liability will equal the expected harm, as well as it can be ascertained *ex ante*. But basing liability on averages (or substituting fines or taxes on emissions, with payments equal to expected harm, for tort liability) would induce the same behavior.

Accurate information about *ex post* consequences is useful in two types of settings. One, the focus of Section A, is when an individual knows (or might be induced to learn) *ex ante* the actual harm her act will cause. If precise prediction is possible for particular acts, behavior will be better if damages reflect actual harm rather than average harm for a class.

Accurate information about *ex post* consequences is also valuable when an adjudicator has poor information on average harm (in particular, worse information than the actor). Suppose, for example, that there are two groups of actors: the first type causes, on average, a low level of harm (5), and the second type causes, on average, a high level of harm (15).⁶⁰ One option would be to provide liability equal to average harm for both types taken together (10). But then the first type will be overdeterred and the second underdeterred. Another option is to consider average harm, determining averages for each type (so liability is higher for the second type—15 rather than the two groups' average of 10). Suppose, however, that this is impossible (or too costly). A third option is to award damages equal to actual harm. This option might be best, even though individuals of both types could not predict actual harm. The reason is that, under this rule, the first type of individual—who causes, on average, a low level of harm (5)—has a low expected liability (5) and the second a high expected liability (15). It is, of course, possible that in a particular case the low-harm-causing type would pay a high award, reflecting an atypically high level of harm, and the high-harm-causing type would pay a low award, but on average this would not be true. Even though the harm in the particular case is only a single data point on individuals'

⁶⁰ A similar argument would apply if actors faced the same options but one dimension of their actions affected the expected level of harm. (For example, greater care may reduce the expected harm caused by the activity, as modeled in Spier, *supra* note 15.)

types, it is an unbiased estimate, so using it would lead, *ex ante*, to appropriate behavior.⁶¹

The conclusion of Section A must, therefore, be qualified: there are instances in which accurate assessments that cannot be anticipated will be useful. Note, however, that the qualification is modest. For, while individuals cannot predict the actual accurate assessments, they are assumed to be able to predict the distribution of such assessments. Moreover, in this instance the argument depended on the fact that, with accurate assessments, different types of individuals would know that their expected damage payments would, accordingly, differ. If individuals did not know whether they were the type that causes higher or lower harm on average, there would be no value derived from the court determining actual damages *ex post* rather than relying on the average for all such individuals.

4. Scheduling Sanctions

a. *Personal Injuries.* A conscious decision to reduce accuracy by setting damages equal to average harm rather than particular harm is reflected in proposals to provide a damage schedule for personal injuries.⁶² Workers' compensation schemes use such an approach to some extent.⁶³ But scheduling of damages is not generally employed in the tort context.

It is useful to compare other, similar settings in which schedules are used. First-party insurance is a good example. Policies often provide particular payments for loss of limb, rather than indicating that, in the event of injury, an inquiry will be made to determine actual losses. Presumably, if more precise compensation were thought to be worth the

⁶¹ Damages equal to actual harm will not be the best estimate one could make of an injurer's expected harm. To illustrate, if low types cause an average harm of 5 and high types an average of 15, and actual harm in a given case were 20, 20 could hardly be the best estimate of the defendant's expected harm. An adjudicator could set damages equal to the best estimate of the defendant's expected harm (using Bayes' rule) rather than setting damages equal to actual harm. Expected liability would be the same whichever method is used. As suggested by the analysis in Subsection A3, this alternative would be preferable if defendants were risk-averse (because there would be less variance in damage awards) but inferior if plaintiffs were risk-averse (because compensation would be less accurate).

⁶² See, for example, 2 American Law Institute, *Enterprise Responsibility for Personal Injury* 217–27 (1991); Bovbjerg, Sloan, & Blumstein, *supra* note 54. Another proposal involves allowing statistical claims profiles to be introduced as evidence or to determine awards without resort to a trial. See Kenneth S. Abraham & Glen O. Robinson, *Aggregative Valuation of Mass Tort Claims*, 53 L. & Contemp. Probs. 137 (1990).

⁶³ See Donald T. DeCarlo & Martin Minkowitz, *Workers Compensation Insurance and Law Practice* 82–83 (1989).

cost of particularized inquiries, policy provisions would differ.⁶⁴ Another instance is disability insurance, where policies typically provide for a fixed percentage of one's prior wage for covered disabilities.⁶⁵

In other first-party insurance contexts, such as homeowners' insurance, there is a mix of scheduling and inexpensive forms of alternative dispute resolution. Scheduling arises implicitly from advance appraisals of particular objects, so it is particularized.⁶⁶ In other instances, or when property is damaged without losing all its value, a common provision is for binding arbitration.⁶⁷ First-party auto insurance for injury by uninsured motorists operates similarly.⁶⁸

Thus, in a wide range of contexts, particularly including contractual settings, one observes damage scheduling or the use of very inexpensive dispute resolution. It is unclear whether these practices should be seen as demonstrating the superiority of employing less accuracy in conventional adjudication. On one hand, the questions posed often are the same. On the other hand, *ex ante* provision by contractual arrangement offers advantages not available when contracts are silent (rather than providing for liquidated damages) or in contexts such as accidents in which there is no prior contractual arrangement.⁶⁹ In addition, the use of alternative

⁶⁴ Health insurance policies, of course, typically reimburse medical expenses, which are particularized. (Policies also typically omit entirely pain and suffering, most probably because it is inefficient for individuals to insure against such nonpecuniary losses, see note 34 *supra*; costs of accurate assessment could be avoided by scheduling.) Life insurance involves scheduling of sorts, in that an insurance policy will provide for a specified payment. But in such instances, as well as with provisions providing a specified amount for loss of limb, the amount is contracted for in advance by the insured party, so it is likely to reflect some particulars of that individual's circumstances. (Nonetheless, there will remain variations that are not taken into account.)

⁶⁵ As discussed previously (see note 64 *supra*), to some extent this reflects that individuals choose the level in advance by contract, to suit their particular situation. But disability coverage tends to be rather standardized (often purchased in a group plan through one's employer that provides the same coverage for all employees who participate), suggesting that the benefits of more precisely tailored compensation are not viewed as worth the costs.

⁶⁶ Nonetheless, the value is usually determined in an inexpensive manner in contrast to disputes over valuation in the context of litigation. This could reflect that precision of compensation is not that valuable. Alternatively, it may be that, with *ex ante* valuations, there is only a small probability that they will ever govern the amount that must be paid, and premiums adjust to reflect expected payments, whereas *ex post* each dollar difference in the valuation will translate directly into a difference in payment.

⁶⁷ See, for example, Thomas W. Mallin & Michael E. Bragg, eds., *Annotations to the Homeowners Policy* 145-46 (2d ed. 1990).

⁶⁸ See, for example, 3 Irvin E. Schermer, *Automobile Liability Insurance* 33-8 to -10 (2d ed., revised 1993).

⁶⁹ Even with accidents, insurance companies might behave in a manner leading to different forms of dispute resolution to the extent these would be advantageous. For example,

dispute resolution may not reflect a differing view about the value of accuracy but, rather, a belief that different procedures can produce as good or better accuracy at lower cost.⁷⁰

b. *Fines*. Fines, as for parking violations, can be formulated in various ways, with different levels of specificity for different settings and with the amount stated *ex ante* or left open to be determined *ex post*. (In the language of Subsection 2, they can be more or less complex, and they may take the form of rules or standards.) Fines often take the form of simple rules (it is the simplicity that is most relevant for present purposes). Thus, there may be a \$10 fine for a parking meter violation, regardless of whether one was illegally parked for ten minutes or ten hours, or there may be a \$15 fine for double-parking, regardless of whether there was no traffic or one caused gridlock for an hour. There are some gradations: blocking an intersection and parking in a taxi space may be treated differently.⁷¹ While further gradation of fines would probably be useful (as the above examples suggest),⁷² no one would advocate any careful inquiry into harm done in particular instances. A number of factors contribute to such a result: the stakes are low, the cost of accu-

an insurance company could state in its policies that, whenever a dispute involved an opposing party with the same clause in its policy, a particular form of alternative dispute resolution would be employed, or both sides would operate in a certain manner in court litigation. Or insurance companies might cooperate informally. (When both parties to a dispute have the same insurer, there generally is no litigation; while determining liability may not be very important in such instances, damages would still matter, as they would affect the payment from the insurance company to the victim, who has first-party insurance.)

⁷⁰ The schemes may use expertise more effectively and avoid duplicative gathering of information and misleading presentation to inexperienced fact-finders. Relatedly, the desire of insurance companies to maintain a good reputation may lead them to provide effective dispute resolution.

Debates on alternative dispute resolution often address accuracy (perhaps indirectly, in calling it second-class justice). See, for example, Edward Brunet, Questioning the Quality of Alternate Dispute Resolution, 62 Tul. L. Rev. 1 (1987). Such debates usually offer little analysis or proof concerning actual differences in accuracy. More relevant for present purposes is that they do not indicate why or how much accuracy is valuable in the particular context.

⁷¹ Many parking tickets list dozens of offenses, although there may be only a few different fine levels. Thus, rather than to determine the penalty, the purpose of the categories is largely to provide information as to the offense—perhaps so the offender will know what the police officer believes was done illegally.

⁷² Further gradation would not be appropriate if enforcement probabilities already reflect differences in expected harm, so that the expected sanction equals expected harm even if sanctions imposed are not proportional to expected harm. (Police may use their discretion in a manner such that tickets are more often given for more serious behavior. But during times when traffic is busy, police may be more heavily engaged in directing traffic and thus have less time to ticket those who double-park near an intersection.)

rately measuring harm may be high, and individuals at the time they park illegally would know only approximately how much disruption they would cause.⁷³

In other regulatory settings, such as the violation of pollution or safety standards, fines also tend not to be particularized. In these instances, another factor may contribute to the desirability of such fines: harm may be probabilistic and not have occurred.⁷⁴ Alternatively, it may be that it is extremely difficult to trace particular injuries to particular violations.⁷⁵ Thus, average expected harm may be far cheaper to ascertain than actual harm. Moreover, there is little difficulty with regard to diminished incentives because the injurer also is unable to determine actual harm at the time she decides how to act.⁷⁶

c. *Criminal Sentencing.* In the past decade, the federal government has adopted criminal sentencing guidelines, involving a highly detailed sanction schedule.⁷⁷ For present purposes, two features of these guidelines are notable. First, they often involve substantial differentiation among offenses causing different levels of harm—providing different sentences, for example, for different degrees of offenses against the person and for different amounts of money involved in crimes such as theft or fraud.⁷⁸ In most instances, one suspects that individuals contemplating such crimes would know at least approximately how severe the resulting harm would be. (For example, a thief usually knows whether a crime is likely to involve a few thousand dollars or a few million.) Thus, making

⁷³ To illustrate, determining actual harm might require knowing who was inconvenienced and what mission was delayed as a result. This inquiry would be extremely expensive, and a violator could not readily anticipate the outcome. In contrast, if one double-parks near an intersection during rush hour, one can anticipate causing significant disruption, and it would be simple for a police officer to ascertain, at least approximately, the significance of the disruption.

⁷⁴ One might ask why there is a fine when there is no harm, rather than liability for harm done if and when harm arises. In addition to the argument that follows in the text, an important reason is that injurers may be judgment-proof for the large harm they might cause (but not for the lower fines for violations detected in advance). See Steven Shavell, *Liability for Harm versus Regulation of Safety*, 13 J. Legal Stud. 357 (1984).

⁷⁵ See *id.*

⁷⁶ This is one of the cases noted in Subsection 3.

⁷⁷ See United States Sentencing Commission, *Guidelines Manual* (1992). The motivation for the guidelines involved in significant part concerns for a more open process and for uniformity (see *id.* at 2), suggesting values of promoting fairness and avoiding potential abuses of power (see Subsection VC3 *infra*) in addition to instrumental objectives of the criminal law, such as effectively combating crime. Reducing the cost of adjudication was not a primary issue, as the previous discretionary system did not devote significant resources to determine an appropriate sentence.

⁷⁸ See *id.* §§ 2A, 2B1.1, 2F1.1.

sanctions depend on harm is desirable with regard to its effect on behavior.

Second, the guidelines do not usually require expensive adjudication of the degree of harm. For example, in determining the amount of the loss for theft and related offenses, the commentary indicates that "[t]he loss need not be determined with precision, and may be inferred from any reasonably reliable information available, including the scope of the operation."⁷⁹ And facts are not determined at a trial but, rather, using less formal sentencing procedures.⁸⁰ Thus, the system operates in a manner suggesting that approximate accuracy is a reasonable objective, while great precision is not worth the additional costs involved. An important exception, however, is that some of the differentiation provided by the guidelines involves different sentences for formally different crimes rather than for different degrees of harm from a given type of crime. Establishing the crime category requires proof beyond a reasonable doubt at trial or a plea bargain made with the knowledge that the prosecution would otherwise have had to meet such a standard.⁸¹

5. Market-Share Liability

When it is difficult to determine which of many potential injurers caused a particular plaintiff's injury, market-share liability has been proposed and adopted in some instances.⁸² For example, if four manufacturers each sold an identical drug that caused delayed illness in some users, each manufacturer might be held responsible for the fraction of damages corresponding to its percentage of sales in the relevant period. Arguments in favor of such an approach usually combine the difficulty of determining

⁷⁹ *Id.* § 2B1.1, commentary, application note 3.

⁸⁰ See *id.* § 6A1. In particular, when a matter is disputed, section 6A1.3 provides for information to be submitted in addition to the presentence report, but the commentary indicates a preference for written submissions of counsel and affidavits when possible. (The commentary further indicates that sentencing judges are not restricted to admissible information and that a preponderance of evidence standard is appropriate, which presumably reduces the prosecution's need to offer expensive proof unless the defendant is able to present credible contrary information on the disputed factor.)

⁸¹ This, of course, raises the question why proof that one has committed a more serious crime is subject to high standards guaranteed by the Constitution while proof of the degree of a given offense—which may have as great an effect on the sanction—is not subject to such requirements.

⁸² See, for example, *Sindell v. Abbott Laboratories*, 26 Cal. 3d 588, 607 P.2d 924, 163 Cal. Rptr. 132, cert. denied, 449 U.S. 912 (1980); David Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 Harv. L. Rev. 849 (1984). A related approach involves the use of statistical sampling of cases to determine damages. See Robert G. Bone, *Statistical Adjudication: Rights, Justice, and Utility in a World of Process Scarcity*, 46 Vand. L. Rev. 561 (1993).

causation, the need to provide incentives for potential injurers to behave properly, and the desire to compensate victims.

Observe that in the context just described, a market-share approach may be preferable *even if causation from particular manufacturers to particular injuries could be established*. First, establishing causation would cost more. Second, establishing causation would not alter the expected liability of each manufacturer, so there is no behavioral gain to be had from accuracy. Third, it would not affect victim compensation in the case posed. Moreover, in some important variations, victim compensation would be worse from an insurance perspective if individualized causation were established. Suppose, for example, that one manufacturer with 25 percent of the market was bankrupt. Then, under market-share liability, each victim would receive 75 percent compensation. If, instead, causation were accurately traced, 75 percent would receive full compensation and 25 percent would receive nothing. Victims as a class are worse off because their situation is riskier, while manufacturers are unaffected.⁸³ Thus, even if causation can be established accurately, it may cost more and produce a less desirable outcome.

There is one obvious and important caveat. The above analysis assumes that it is known that the products or processes that may have caused the injury are fungible.⁸⁴ If this is not known, tracing causation to particular victims would be a way of determining how much harm each manufacturer actually caused. This argument parallels that in Subsection 3 concerning the benefit of determining actual harm when expected harm caused by an individual or group is not known. Note, however, that tracing harm in perhaps thousands of particular cases may not be the best way to gather the relevant information. A statistical study may be of similar accuracy and far less costly.⁸⁵ Even if it is less accurate, it

⁸³ Accurately tracing causation would introduce additional sources of uncertainty. For example, if 50 percent of illness was due to natural causes, a market-share approach (properly conceived) would provide 50 percent compensation to everyone rather than 100 percent compensation to half of those who were ill. Also, accurate tracing in any setting will impose some risk on defendants. One's expected liability has less variance when one is held responsible for one's market share than for the injury one actually caused (which may be more or less, even though it is, on an expected-value basis, precisely equal to one's share).

⁸⁴ It is also assumed that the injuries are caused by the manufacturers rather than other sources, which suggests similar qualifications.

⁸⁵ It could not be more accurate than determining the actual cause of harm in each and every case. Many errors, however, will be made investigating each case, particularly given the methods of proof employed in adjudication. (In particular, victims have incentives to mislead the fact-finder about causation since their compensation depends on it. If, instead, they were asked to participate in a statistical study that would determine manufacturers' liability, but they were provided a level of compensation that was independent of the outcome of the study, the information obtained might be better.) Particularly when illness

may be sufficiently cheaper to justify its use in place of conventional tracing of particular injuries.

C. Parties' Incentives to Present Information in Adjudication

The discussion in Sections *A* and *B* largely assumes that the legal system—those who promulgate rules or judges who apply them—is able to choose the degree of accuracy in adjudication. In adversary adjudication, however, this choice is made largely through a scheme of indirect controls. Legislators and judges do not determine which witnesses will testify at trial. Rather, they establish a system in which private litigants present whatever information they believe will further their cause. The quantity and type of information a fact-finder ultimately considers is thus a product of procedural and substantive law and the parties' choices. This section considers whether parties' incentives to present information tend to be socially appropriate, excessive, or inadequate.⁸⁶

This question has an important bearing on the design of the legal system. If parties' incentives tend to be socially appropriate, there would be no need to worry greatly about too much or too little accuracy. One could leave it to the litigants, who would produce the correct level.⁸⁷ If

has multiple causes, statistical evidence may be the *only* reliable evidence available in any event. Relatedly, in statistical studies with large samples, a fairly high error rate in each observation may be unimportant as long as errors in measurement are unbiased because errors will tend to cancel. Adjudication of an individuals' right to recover is not usually conducted under the assumption that error is permissible because it evens out across cases. See Section *C* *infra*.

⁸⁶ As noted in the Introduction, questions of whether a given system produces accuracy at the least cost—for example, whether an adversary system is an efficient producer of information—are beyond the scope of this inquiry. Also, the reader should keep in mind that the present analysis concerns incentives to present factual information rather than incentives to develop and present arguments about the appropriate content of the law. While many elements of the analysis here are applicable in the latter context, there is the additional consideration that determination of the law in a particular case is a public good to the extent the case serves as a precedent, a factor suggesting that private incentives would be insufficient. See also Subsection *B2c* *supra*.

⁸⁷ This is a familiar invisible-hand argument. Note that a general belief in the efficient functioning of markets does not provide any basis for inferring good results in adjudication (unless the form of adjudication was itself chosen in the market, as when it is specified by contract). The reason is that virtually every act of a litigant, by design, hurts the opponent; externalities are thus a central characteristic of behavior in litigation. This fundamental difference between litigation and other goods and services is often overlooked. See, for example, Jerry L. Mashaw, *Bureaucratic Justice: Managing Social Security Disability Claims* 80–81 (1983) (arguing that litigation, unlike dispute resolution concerning benefit determinations, is largely analogous to other markets, such as that for cars, with respect to individuals' incentives to incur litigation costs); Jonathan R. Macey, *Rule of Law: Not All Pro Bono Work Helps the Poor*, *Wall St. J.*, Dec. 30, 1992, at A7 ("When clients must pay for legal services, lawyers will be hired only when the benefits of the lawyers' activities outweigh the costs.").

parties' incentives might be incorrect, then one must take this into account in designing adjudication. Alternatively, one may view the question of the appropriate degree of accuracy as involving, at the implementation phase, the design of a set of procedures so that self-interested parties will be induced to produce the right amount of accuracy (at the least cost).

1. Analysis of Parties' Incentives⁸⁸

Return to the scenario from Section A, in which a damage award may be based on the average harm for acts in the class or the actual harm caused by a particular act. It was assumed that the legal system chose whether or not more accurate information would be used. Now assume instead that the adjudicator begins with information about the average harm; damages will equal average harm unless one of the parties presents reliable evidence of actual harm, in which case damages will equal actual harm.⁸⁹

When will a plaintiff or defendant offer evidence of actual harm? For concreteness, assume that average harm is 150. Furthermore, assume that, after the accident, each of the parties learns of the actual harm and can prove it in adjudication at a cost of 20. Then, whenever actual harm is less than 130, the defendant will spend 20 to prove actual harm; whenever actual harm exceeds 170, the plaintiff will spend 20 to prove harm; and whenever actual harm is between 130 and 170, neither party will prove harm, so damages will be 150. (The reasoning is straightforward: if actual harm were 125, the defendant would rather pay 125 in damages plus 20 in litigation costs than pay 150 in damages. If actual harm were 175, the plaintiff would rather collect this amount, spending 20 to prove it—for a net of 155—than collect damages of 150. If actual harm is near 150, neither party gains enough by proving that actual harm differs from 150 to justify the cost of 20.)

The remainder of the analysis in this subsection explains why the incentives just described are not generally socially appropriate. The discussion first considers the case in which individuals are uninformed of actual harm *ex ante*, then the case in which they are informed, and finally concludes with remarks on some of the simplifying assumptions employed in the analysis.

⁸⁸ For a formal analysis of most of the argument in this subsection, see Kaplow & Shavell, *Accuracy and Damages*, *supra* note 5, at 11–17.

⁸⁹ This description is simpler than actual adjudication in that it ignores the possibility of conflicting, misleading, or unreliable evidence. In some sense, it is an optimistic scenario for the presentation of information. It assumes that each party is able, with certainty, to establish the truth at some cost (borne entirely by itself).

a. *Whether Incentives to Present Information Are Excessive When Individuals Cannot Anticipate the Actual Level of Harm.* When injurers do not know at the time they act what the actual harm will be, expenditures ex post by plaintiffs and defendants to prove the actual harm are a pure waste. The analysis in Subsection A1a explained that greater accuracy is of no value because it cannot improve injurers' behavior ex ante.⁹⁰ Yet, in adjudication, ex post, both injurers and victims will sometimes find it in their interest to make expenditures establishing actual harm, if such evidence is admissible.

Two sorts of policies might address this problem. One could adopt a rule that ignored proof of actual harm.⁹¹ Alternatively, some sort of tax on presenting information could be implemented.⁹² Such a tax would raise parties' private costs of presenting information without consuming real resources.⁹³

b. *Whether Incentives to Present Information Are Excessive When Individuals Do Anticipate the Actual Level of Harm.*⁹⁴ When injurers know the actual harm they will cause at the time they act, the prospect of greater accuracy ex post as a result of parties' incentives in litigation will improve behavior. The question is whether the value of the improvement is sufficient to justify the ex post cost. In addition, there is the possibility that there may be valuable but unrealized improvements in behavior because parties, ex post, would not have a sufficient incentive to present information leading to accurate results.

The analysis, which is substantially more complicated than for the pre-

⁹⁰ There might be a benefit of accuracy with regard to guiding future actors' behavior if the outcome of an accurate adjudication is effectively disseminated. See Subsection B2c *supra* (indicating that often this is unlikely).

⁹¹ In Subsection b, where it is assumed that individuals know actual harm ex ante and thus information has some value but private incentives may be excessive, the analogue would be to ignore proof of actual harm when it differs only modestly from average harm. Ruling that evidence is inadmissible because it is largely redundant of other evidence has this feature. A party may wish to present evidence at a cost of 10 that will affect the outcome by an expected amount of 11. Even if the party's evidence is entirely reliable and will not lead to further expenditures by the other side, such evidence may have a cost in excess of its social value.

⁹² One could imagine options such as taxing legal services in litigation or charging fees for court time.

⁹³ One can contrast legal rules that make it more difficult to present information. If the information is nonetheless presented, more real resources will have been consumed, whereas a tax transfers money without wasting resources.

⁹⁴ The effects examined in this subsection are of the sort identified with regard to incentives to file suit in Steven Shavell, *The Social versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. Legal Stud. 333 (1982).

ceding case, is sketched in the margin.⁹⁵ The basic points are as follows. Initially, it can be demonstrated that the social benefit from improving behavior is generally less than the difference between actual harm and average harm. In contrast, the private benefit from proving harm in adjudication equals the difference between actual and average harm, as explained previously. This suggests that, even when individuals are perfectly informed of actual harm at the time they act, litigants' incentives to present information about harm in adjudication are never inadequate and are sometimes socially excessive. Indeed, this is the case for defendants.

⁹⁵ Suppose that actual harm is 170 rather than the average value, 150. If damages will be 170 rather than 150, how will this affect behavior? For individuals whose benefit from committing the act exceeds 170, there will be no effect: they will act regardless. (This example assumes that harm is certain. If harm has a probability of, say, .1, one could consider individuals with a benefit of 17 rather than 170, 15 rather than 150, and so on, and the analysis would be the same; the maximum benefit of information would be the difference between actual and average harm, weighted by the probability of harm.) The expenditure of 20 ex post will be a waste with respect to them. For those whose benefit from committing the act is less than 150, there will be no effect: they will not act in either instance. (There is no waste, however, because there will never be an occasion ex post to spend the 20 to prove that harm is 170 rather than 150.) For one with a benefit of 160, the act will be committed if damages are 150 but not if they are 170. The social gain from deterring this act is 10 (the difference between the harm of 170 and the benefit of 160). For one whose benefit is 151, the social gain from deterrence will be 19. For one with a benefit of 150.01, the gain will be 19.99. Thus, 20—the difference between the actual harm and average harm—is the maximum possible benefit from accuracy. In virtually all cases (all but those where the private benefit exactly equals 150 and the individual, indifferent between acting and not acting when damages are 150, chooses to act), the benefit is less than 20, and in many (when the benefit exceeds 170), the benefit is zero.

Return now to the incentive to present information. It was explained that when actual harm is between 130 and 170, neither party would spend 20 to present information. This is desirable with respect to defendants, for information costs 20, while 20 is the maximum possible benefit with regard to behavior when harm is in this range. For plaintiffs, this result is desirable if the act would be committed in any event (for then there is no benefit from demonstrating that actual harm exceeds average harm). But if the act would have been deterred—for example, if actual harm is 160 and the actor's benefit is 155—it would be desirable for harm to be demonstrated, for the prospect of such demonstration provides deterrence, in which case the ex post information cost of 20 is never borne.

Consider now the situation in which actual harm is just below 130 or just above 170. In this case, parties would present information, at a cost of 20. This result usually would be undesirable for demonstration by defendants. The reason is that the maximum possible benefit of the information is close to 20, suggesting that the average benefit of information (which usually is less than 20 and may well be zero) is less than 20. Since the information costs 20 to present, the incentive often is excessive. For plaintiffs, the above reasoning applies: the result is undesirable if behavior is unaffected but desirable if there is deterrence because then the 20 need never be spent.

Finally, consider situations in which harm is much less than 130 or much more than 170. Then it is possible that the benefit of establishing actual harm will, on average, exceed 20 for both plaintiffs and defendants. See also Kaplow & Shavell, *Accuracy and Damages*, *supra* note 5, at 14–17 (formally analyzing the case where the behavioral benefit involves the actor's level of care). The preceding analysis assumes that parties bear their own legal

When plaintiffs have an incentive to demonstrate harm, however, there is a further consideration: the prospect that they will demonstrate an above average harm will sometimes deter harmful acts, in which case the cost of demonstrating harm ex post will not be borne. In such instances, the improvement in behavior is obtained for free, so the ex post incentive is not excessive and may be inadequate. Thus, when actors are fully informed ex ante about the actual harm caused by their acts, plaintiffs' incentives ex post may be excessive or inadequate, depending on whether the particular act would in fact be deterred by the expectation that actual harm would be established in adjudication.

c. *Comments on Simplifying Assumptions.* The preceding discussion indicates that defendants' incentives to present information are systematically excessive, while plaintiffs' are excessive when actors do not know actual harm when they act and indeterminate when actors are informed when they act. The analysis oversimplifies in many respects. First, it ignores risk aversion. As noted in Subsection A3, when plaintiffs are risk-averse, accuracy may be more valuable than otherwise and, when defendants are risk-averse, accuracy may be less valuable. Thus, for example, if plaintiff risk aversion were a significant problem (because plaintiffs were very risk-averse and could not obtain insurance), accuracy would be more valuable, so what appeared to be an excessive incentive to present information in adjudication may not be so excessive.⁹⁶

Another simplification is that the possibility of settlement is ignored. When settlement is possible, accuracy may be less costly. In the extreme case in which settlement was certain before any legal costs were incurred and settlement reflected actual harm rather than average harm, accuracy would be free. Thus, if accuracy has any positive value, it would be desirable. More realistically, settlement often occurs after substantial expenditures, the magnitude of which is related to the level of accuracy.⁹⁷

costs. If plaintiffs' costs were borne by defendants, as would be efficient in this scenario with regard to effects on injurers' ex ante behavior, plaintiffs' incentives to demonstrate harm would unambiguously be excessive.

⁹⁶ When individuals are uninformed of actual harm ex ante, the ex post incentive is still socially excessive. Because there is no benefit with regard to improving behavior, the only possible benefit would be enhanced accuracy in compensation. But since parties are willing to spend up to a dollar for each dollar change in the outcome, expenditures would be excessive even if plaintiffs were risk-averse. (Even a risk-averse individual would not be willing to spend up to a dollar to improve the precision of his insurance coverage by a dollar.) See Subsection IVA2 *infra*.

⁹⁷ For an analysis of how legal complexity (similar to greater accuracy; see Subsection B2a *supra*) affects the likelihood of settlement when there is asymmetric information concerning actual harm, see Spier, *supra* note 15. Also note that accuracy, by making trial more expensive, may make settlement more likely, which could reduce litigation costs. That settlement is ignored in Sections A and B is of little consequence, for the analysis

For example, if particular aspects of the plaintiff's circumstances are relevant in establishing damages, then investigation, taking of depositions, and consultation of experts may have occurred before settlement is reached. It is even possible that costs in proving actual harm could be greater than suggested in the above example because, even if costs are lower in each case, they may be incurred more often when bargaining over settlement amounts is taken into account.⁹⁸

In conclusion, whether private parties' incentives to provide information in adjudication are socially appropriate is a complex question. When injurers are uninformed about actual harm *ex ante*, the *ex post* incentive is clearly excessive; when injurers are perfectly informed, the incentive may well be excessive, although not always, in the simple case, while settlement and other values of accuracy further complicate the calculus.⁹⁹ This subsection also employed other simplifying assumptions, such as that parties who spend resources to prove actual harm do so only when what they are proving is the truth and that they will be successful in proving the truth. Overall, it seems quite clear that there is no strong basis for assuming that parties' incentives generally tend to be appropriate.¹⁰⁰

there simply assumed that greater accuracy involves some added cost. As long as settlement does not always occur (and the frequency is not greatly affected by accuracy) or involves some added cost when adjudication is more accurate, the analysis would be qualitatively the same. (As noted in the Introduction, if accuracy reduced costs overall, one could usually conclude that it was desirable without bothering to analyze its benefits.)

⁹⁸ For example, if a plaintiff has injuries that are above average, this may not be known to the defendant, who then would not be willing to settle for a high amount. Thus, there must be enough gathering and exchange of information for the defendant to believe that the plaintiff's likely recovery is high in order for the plaintiff to obtain a generous settlement. With such asymmetric information, there may tend to be an unraveling phenomenon to some extent. See Bruce L. Hay, *Civil Discovery: Its Effects and Optimal Scope*, in this issue; Steven Shavell, *Sharing of Information prior to Settlement or Litigation*, 20 *RAND J. Econ.* 183 (1989). Plaintiffs with above average harm will prove this to defendants. This leads defendants to revise downward their estimate of the average harm of plaintiffs who remain silent. Then, of this group, those with greater harm will have an incentive to prove their actual damages, and so on. Thus, it is possible that the cost of establishing harm would be incurred in most cases, even cases involving actual harm close to or below average harm: costs may not be incurred at trial, but costs would be incurred before trial, in the process of bargaining over settlement amounts.

⁹⁹ If parties were uninformed *ex ante*, but could make expenditures to learn more, the analysis would involve aspects of both cases, as well as additional complications concerning *ex ante* decisions to acquire information.

¹⁰⁰ See note 87 *supra* (no reason to suppose that the invisible hand produces efficient expenditures on litigation because litigation involves imposing costs on opposing parties). Note that the social value of accuracy will vary greatly by context, while the *ex post* incentives to present information tend to be similar and largely independent of the social value of accuracy.

2. Applications

Growing concern with the cost of litigation has led to reform proposals designed to reduce costs.¹⁰¹ Some reforms are designed to reduce the presentation of information, such as limits on the amount of discovery or the numbers of expert witnesses.¹⁰² In addition, judges often exert informal, but powerful, pressure on parties to act in ways that reduce costs of disputes.¹⁰³ The analysis here provides a framework for evaluating such proposals and actions.

To the extent that incentives to present information in litigation are more often excessive than inadequate, the analysis also casts doubt on the wisdom of rules that require parties to present more information than they would present on their own. Some evidence rules—those excluding hearsay,¹⁰⁴ requiring “best evidence,”¹⁰⁵ demanding that foundations or chain of custody be established¹⁰⁶—force parties to present more accurate but more expensive information than they would choose. Production burdens of proof operate similarly. There may be justifications for many such rules because they may reduce costs if they dispose of cases or eliminate channels of proof.¹⁰⁷ There also may be concerns about jury error.¹⁰⁸ Nonetheless, it is interesting that some components of the legal

¹⁰¹ While there is much noise about the problem of excessive litigation costs, there is little analysis indicating how we know that the costs are excessive, rather than simply higher than we would “like.” (Presumably, we would like them to be zero.) In particular, the social purposes of litigation usually remain unstated. It is not possible to analyze whether expenditures are excessive when there is no benchmark for evaluation.

¹⁰² See, for example, Judicial Conference, Standing Committee on Rules of Practice and Procedure, Proposed Amendments to the Federal Rules of Civil Procedure and Forms, Rule 16(c)(15) (April 1993; effective December 1993) (authorizing pretrial orders limiting time allowed for presentation at trial); *id.* Rule 30(a)(2)(A) (limiting number of depositions that may be taken without leave of court); *id.* Rule 33(a) (limiting interrogatories available without leave of court); Wall St. J., June 30, 1992, at B7 (reform of discovery and limits on expert witnesses adopted in Arizona).

¹⁰³ At the same time, they regularly fail to rule on motions in a timely fashion, leading parties to spend on discovery vast sums that turn out to be wasted even from a private point of view.

¹⁰⁴ Fed. R. Evid. 802.

¹⁰⁵ Fed. R. Evid. 1002.

¹⁰⁶ An example is the business records exception to the hearsay rule, Fed. R. Evid. 803(6), although it should be noted that the requirements are much more lenient than at common law. See Advisory Comm. Notes, Fed. R. Evid. 803(6).

¹⁰⁷ Production burdens allow more frequent summary disposition of cases. A hearsay rule may reduce costs if, for example, the declarant is unavailable (assuming that the party is not thereby induced to offer additional, more costly evidence instead).

¹⁰⁸ This concern may explain the hearsay rules—particularly outside the criminal context—although plausible views of jury error provide a weak foundation, and one inconsistent with the many areas in which jurors are trusted. See Note, The Theoretical Foundation of the Hearsay Rules, 93 Harv. L. Rev. 1786 (1980).

system operate to push parties to present more information, while others only occasionally ignore relevant information,¹⁰⁹ even though it is plausible in many contexts that private parties' incentives to present information are excessive rather than inadequate.

III. ACCURACY IN THE DETERMINATION OF LIABILITY

In this part, attention shifts from accuracy in assessing damages to accuracy in determining liability. As noted in the Introduction, errors in adjudication with regard to liability are of two types: failing to impose liability on those who violate legal commands (false negatives), and imposing liability on those who did not violate legal commands (false positives). For convenience, those who violate the law are referred to as the "guilty" and those who do not as the "innocent," and the language of "sanctions" and "penalties" will be employed, even though most of the analysis is applicable to civil and criminal disputes (as well as violations of rules internal to organizations).¹¹⁰ Mistakes in determining liability can arise in many ways. There may be uncertainty concerning the identity of the person who committed an act, whether an act was committed, whether an act in fact caused the victim's injury, or whether an act was justified in some manner recognized by the law.

The analysis in Section *A* emphasizes the relationship between efforts to enhance accuracy and other aspects of enforcement, notably the general level of enforcement effort (audit or investigation rates, intensity of police patrols, rate of private suits) and the level of sanctions. Thus, the analysis will emphasize trade-offs among various enforcement instruments. In addition, the discussion will focus on the costs of imposing sanctions, which received only brief attention in Part II.¹¹¹ Section *B* addresses the relationship between setting the burden of proof and choos-

¹⁰⁹ Notable instances involve excluding evidence on account of prejudicial effects (see Fed. R. Evid. 403) or because of purposes unrelated to accuracy or cost (as with the constitutional exclusionary rules; see note 138 *infra*), rather than out of a concern for cost.

¹¹⁰ Subsection C1 explains how even portions of the analysis that seem uniquely applicable to the criminal context have important implications in the civil context or that of private dispute resolution.

¹¹¹ In particular, Subsection II A3 discussed risk aversion. Imposing risk is an important social cost of sanctions, distinguished from the aspect of monetary sanctions that merely involves transferring funds between individuals or from individuals to the government.

Part II devoted less attention to sanction costs in part because they are less directly affected by policies concerning accuracy in the context of assessing damages. Setting sanctions equal to average harm for a class rather than equal to actual harm for each act in a class involves, as a first approximation, imposing the same amount of sanctions in total. (The approximation is not precise because behavior may differ under the two approaches, which may affect total sanctions. In addition, when sanctions are costly, it is no longer

ing a level of accuracy. Applications of the analysis are presented in Section *C*. Section *D* examines issues involving the extent to which individuals are informed about the law and their incentives to present information in adjudication.

*A. Accuracy, Enforcement Effort, and Sanctions*¹¹²

Consider a scenario in which individuals decide whether to commit an act that causes a known level of harm. (Such an act might be a crime, tort, breach of contract, violation of a disclosure obligation, or whatever.) It is assumed that individuals commit acts when their benefit from so doing exceeds the expected sanction cost.¹¹³

The legal system provides a sanction for such acts. (The sanction may be a fine, imprisonment, or damages paid to the victim.) For present purposes, the level of the sanction is not taken as given but is viewed as involving a choice in designing the system.¹¹⁴ Raising the sanction reduces

optimal to set expected sanctions equal to expected harm, which would further complicate the analysis. See, for example, A. Mitchell Polinsky & Steven Shavell, *The Optimal Use of Fines and Imprisonment*, 24 J. Pub. Econ. 89 (1984).) Thus, that sanctions are costly need not be relevant to the choice. Risk aversion is an exception because risk depends not only on the aggregate amount of sanctions imposed but, for defendants, on their variance and, for plaintiffs, on the extent to which they reflect actual harm. (In contrast, costs of imprisonment, as a first approximation, depend on the total amount of imprisonment: two five-year terms and one ten-year term result in about the same costs of maintaining prisons and involve the same total deprivation of liberty. Restated formally, costs of nonmonetary criminal sanctions may be approximately linear.)

¹¹² Most of the analysis in this section is demonstrated in a formal model in Louis Kaplow & Steven Shavell, *Accuracy in the Determination of Liability*, 37 J. Law & Econ. (in press, April 1994). The most relevant prior economic literature concerns incentives in principal-agent relationships; some of that literature discusses the value of information in monitoring. See Sanford J. Grossman & Oliver D. Hart, *An Analysis of the Principal-Agent Problem*, 51 *Econometrica* 7, 35–38 (1983); Bengt Holmström, *Moral Hazard and Observability*, 10 *Bell J. Econ.* 74, 81–88 (1979); Steven Shavell, *Risk Sharing and Incentives in the Principal and Agent Relationship*, 10 *Bell J. Econ.* 55, 64–65 (1979).

¹¹³ Both elements of this statement may be interpreted broadly. Their benefit may be defined net of any aversion toward violating the law for its own sake (that is, without regard to the risk of sanctions). The expected sanction cost in some instances will refer not to the cost of some expected sanction (the mean) but rather to whatever cost one associates with the possibility of sanctions (thus allowing for the possibility of risk aversion). The discussion will proceed on the assumption that individuals' benefits vary, so that, for any given legal regime, some individuals will be deterred and others will act. Also, the language is that of acting or not acting, but the decision may be taken to include how one acts (so that acting may be acting without taking precautions and not acting may be not acting as dangerously, by taking precautions); relatedly, "innocent" individuals may be individuals who acted in a less harmful manner rather than in a harmless manner.

¹¹⁴ In the civil context, the level of damages is affected by various rules determining which elements of harm are compensable (consequential damages, pain and suffering) and

the number of individuals who will commit the act (assuming other aspects of the system remain unchanged). This section considers the case in which sanctions are socially costless (as with fines or civil damages, that are mere transfers if individuals are not risk-averse) and the case in which sanctions are socially costly (as with imprisonment, or monetary sanctions when individuals are risk-averse).

The legal system also involves a level of enforcement. Thus, for tax compliance, there is an audit rate; for reported crimes, an investigation rate; for traffic violations, a density of police on patrol; for civil disputes, a set of rules affecting the likelihood that a victim will sue. Increasing enforcement effort deters more individuals. Such increased effort is costly.¹¹⁵

Finally, designing the legal system entails choosing a level of accuracy. The greater the level of accuracy, the fewer innocent individuals are sanctioned and the more guilty are sanctioned. One can view enforcement effort as determining the number of individuals who are detected and sanctioned, while accuracy indicates the fraction of those sanctioned who are guilty.¹¹⁶ For example, one can choose the number of audits and the care with which each is conducted. Or one may determine how many thefts to attempt to solve and how carefully to adjudicate the guilt of each suspect. Greater accuracy is assumed to be costly.

This section will analyze three ways in which accuracy in the determination of liability may be valuable. First, greater accuracy is a means of achieving deterrence,¹¹⁷ in addition to raising sanctions or enforcement

whether damages simply equal harm or may differ (as with damage multipliers, statutory minimum damages, punitive damages). In private arrangements, sanctions may be specified by contract (liquidated damages), determined by an arbitrator, or left to the default rules provided by the formal legal system.

¹¹⁵ A qualification is that raising enforcement rates in some contexts may reduce enforcement expenditures on account of the increase in deterrence. For example, if a sufficiently high percentage of cases were pursued vigorously, few individuals would continue to commit such acts. (In contrast, posting more officers along a highway may reduce speeding, but that reduction will not in turn reduce the number of officers who must be posted.) Most of the discussion to follow, however, will consider different combinations of techniques that achieve a given level of deterrence.

¹¹⁶ Accuracy and enforcement effort are not inherently separable. For example, hiring better detectives may result in more suspects being detected, with a higher fraction of those detected being truly guilty. Nonetheless, it is useful in thinking about the problem to view each aspect separately.

¹¹⁷ "Deterrence" will often be used as a synonym for controlling behavior. Although the term is more familiar in the criminal context, where for most crimes it would be ideal to deter all harmful acts (if this could be done at no cost), the term can be applied to typical cases in the civil context. For example, increasing deterrence for an activity subject to tort law (that is often desirable when conducted properly, but not always, at least when con-

effort. Second, when deterrence is achieved through enhanced accuracy rather than by using alternatives, sanctions are imposed less often, which is a benefit to the extent that sanctions are socially costly. Third, increasing accuracy may increase the precision with which behavior is controlled.

1. Accuracy and Deterrence

Accuracy is relevant in controlling behavior because increasing accuracy, like increasing the level of sanctions or enforcement effort, is a method of increasing deterrence.¹¹⁸ The reasoning involves two components.¹¹⁹ First, greater accuracy—holding sanctions and enforcement effort constant—increases the likelihood that the guilty are sanctioned rather than mistakenly exonerated. Thus, individuals contemplating whether to act expect the likelihood of sanctions to be higher if they commit the harmful act. (The likelihood is the product of the probability that they will be detected and the probability that they will be sanctioned given detection. Increasing enforcement effort raises the first factor and increasing accuracy raises the second factor.) Second, greater accuracy reduces the likelihood that the truly innocent are sanctioned. This makes a decision not to commit the act look more attractive. Both factors—making harmful acts less attractive and harmless behavior more attractive—increase deterrence.¹²⁰

It might be objected that many sources of enhanced accuracy do not

ducted improperly) simply means increasing the expected liability associated with that activity.

¹¹⁸ Prior literature has emphasized how the presence of legal error may affect deterrence and, in some instances, how adjusting sanctions may be an appropriate response. See, for example, Isaac Ehrlich, *The Optimum Enforcement of Laws and the Concept of Justice: A Positive Analysis*, 2 *Int'l Rev. L. & Econ.* 3, 16–18 (1982); I. P. L. Png, *Optimal Subsidies and Damages in the Presence of Judicial Error*, 6 *Int'l Rev. L. & Econ.* 101 (1986); A. Mitchell Polinsky & Steven Shavell, *Legal Error, Litigation, and the Incentive to Obey the Law*, 5 *J. L. Econ. & Org.* 99 (1989); Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 *J. Legal Stud.* 399, 402–10 (1973). Such literature has not, however, considered the level of accuracy as an instrument of enforcement policy and, thus, has not considered the value of accuracy in determining liability.

¹¹⁹ Accuracy may also influence individuals' efforts to evade sanctions. Such efforts may be reduced if greater accuracy renders them ineffective or increased by inducing shifts to more costly techniques that tend to be successful in spite of more accurate procedures.

¹²⁰ If the likelihood of mistakenly sanctioning the innocent is small, as one often suspects is the case, this second factor would be much less important than the first. Sometimes, however, it may be important. Many areas of the law are ambiguous or difficult to apply without a significant risk of error (such as some tax rules). Then, individuals might reason that committing an act is beneficial in significant part because they are fairly likely to be sanctioned even if they do not commit the act. Moreover, which situation prevails depends on how the burden of proof is set. See Subsection *B2a infra*.

reduce both the likelihood of mistakenly exonerating the guilty and of mistakenly sanctioning the innocent. In fact, some—such as providing free defense counsel to indigent criminal defendants—may increase one type of error (mistaken acquittals) while reducing the other (mistaken convictions). As Section *B* will emphasize, however, such an argument involves changing both accuracy and (implicitly) the burden of proof. If the *de facto* burden of proof is held constant (in a manner to be explained), increased accuracy will by definition result in reductions in both types of error.

Having described how accuracy, enforcement effort, and sanctions affect deterrence, it is now possible to explore what combination of these three aspects of enforcement policy should be selected.¹²¹ The simplest answer, reflecting the economic perspective of this article, is to choose the combination that minimizes costs. The remainder of this subsection examines factors relevant in determining the appropriate combination.

a. *Sanctions and Enforcement Effort.* First, consider sanctions and their relationship with enforcement effort. If sanctions are costless (as when they involve mere monetary transfers and individuals are risk-neutral), there is a benefit in employing them to the maximum feasible extent. If the sanction were less than its maximum feasible level, it could be raised, thereby increasing deterrence at no cost. Then, one could reduce enforcement effort by just the amount that restores deterrence to its prior level. The combination of these two changes has the result that the control of behavior remains unaffected but enforcement costs are reduced.¹²² Therefore, as long as the sanction is less than its highest possible level, the enforcement policy is not optimal.

¹²¹ The emphasis in this part is on achieving a specified level of deterrence in the most appropriate manner. There is, of course, the separate question of how much deterrence is optimal. (It is separate because, whatever level is optimal, it is best to achieve it in the most efficient manner.) When enforcement is costly, the answer is complicated. The simple maxim that expected sanctions should equal expected harm, so that acts will be committed if and only if their benefits exceed their costs, does not hold when there is a cost of obtaining that result. See, for example, Polinsky & Shavell, *supra* note 111. Moreover, the cost of achieving an additional increment to deterrence will depend on how it is achieved—that is, by what combination of increased sanctions, enforcement effort, and accuracy. Thus, the question of optimal deterrence is hardly independent from the present analysis but, rather, will depend in part on its outcome.

¹²² This argument derives from the reasoning in Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. Pol. Econ. 169 (1968). Ehrlich presents a model in which he argues that higher enforcement with a lower sanction may be optimal. See Ehrlich, *supra* note 118, at 16–18. But he simply assumes that raising enforcement effort does not result in an increase in erroneous convictions. (In addition, his objective function and assumptions about behavior are *ad hoc*—for example, expected sanctions do not determine behavior in any direct way—and there are some errors in his derivations—terms omitted from derivatives and a failure to optimize one of the choice variables.)

The problem of inaccuracy does not fundamentally alter this argument. Increasing sanctions will have less effect on deterrence when there is error because some portion of the increase will fall on the innocent rather than the guilty. But the logic suggesting that it is better to raise sanctions and reduce enforcement effort to save costs still holds.

The literature on the economics of law enforcement has developed numerous qualifications to this result, most of which have similar relevance when one allows for the possibility of error and expenditures to reduce it.¹²³ Of particular importance is the effect of costly sanctions on the analysis, considered in Subsection 2.

Finally, note that, for a given level of accuracy, raising sanctions and reducing enforcement effort need not involve any change in the total extent to which errors are made in imposing sanctions. Thus, for the innocent, mistakenly imposed sanctions will be larger, but they will be imposed less often. These effects tend to be offsetting. (When individuals are risk-neutral, they are precisely offsetting.)¹²⁴

b. Accuracy and Enforcement Effort. Second, consider the appropriate mix of enforcement effort and accuracy. One suspects that attempts to increase each are subject to diminishing returns. For enforcement effort, returns tend to diminish because the best enforcement opportunities are pursued first.¹²⁵ Similar logic holds for accuracy. One might adopt breathalyzers before blood tests if the former achieve most

¹²³ These include differences in individuals' wealth (and thus in the highest fine or damage award they can pay), see A. Mitchell Polinsky & Steven Shavell, A Note on Optimal Fines When Wealth Varies among Individuals, 81 Am. Econ. Rev. 618 (1991); risk aversion, discussed in Subsection 2, see A. Mitchell Polinsky & Steven Shavell, The Optimal Tradeoff between the Probability and Magnitude of Fines, 69 Am. Econ. Rev. 880 (1979) (hereinafter Risk Aversion); limits on the ability to adjust enforcement effort because effort may affect the probability of detection for different types of acts, see Steven Shavell, Specific versus General Enforcement of Law, 99 J. Pol. Econ. 1088 (1991); and individuals' misperceptions of the probability of apprehension, see Lucian Arye Bebchuk & Louis Kaplow, Optimal Sanctions When Individuals Are Imperfectly Informed about the Probability of Apprehension, 21 J. Legal Stud. 365 (1992).

¹²⁴ When individuals are risk-averse, the total imposition of sanctions on the innocent would fall when sanctions are raised and enforcement effort reduced in an amount that keeps deterrence constant because each increment to the sanction has an increasing marginal effect on deterrence. But the utility cost of the higher sanction to the innocent, which seems more relevant, would tend to be constant.

¹²⁵ For investigation of theft, some cases will, by chance, be easier to investigate than others (for example, witnesses vary in their ability to give a precise description). The easiest cases can be pursued first; if a higher rate of detection is required, progressively more difficult cases must be investigated. The same is true with audits if they are not random but, rather, target first the most likely violators. With truly random audits (including inspections and the like), there may be no diminishing returns (except to the extent one must pay progressively higher wages to hire enough auditors, which is unlikely to be a significant problem in the long run).

of the possible improvement over simple observation (smelling breath, having the driver attempt to walk in a straight line) at far lower cost. If more experts improve accuracy in adjudication, one suspects that the twentieth expert does not contribute as much as the tenth.¹²⁶

When two options are subject to diminishing returns, it is often best to employ an intermediate combination. (If one is used substantially and the other not, the former will exhibit little return at the margin while the latter will have a high marginal return, so reducing the first and increasing the second will be more effective.) Similarly, if more deterrence were required, it would seem best to accomplish this by increasing both enforcement effort and accuracy.

There is, however, an important caveat to this logic. The marginal costs of enforcement effort and accuracy are closely interrelated: the higher the level of one, the more costly is the other at the margin. For example, if the criminal courts are made more accurate at a cost of 100 per case, it will be more costly to increase enforcement effort than previously: each added individual who is detected will now be processed through a system that costs 100 more. Or, if the audit rate is high, the total cost of increasing the accuracy of audits will be higher than if the audit rate is low because the increased cost of each audit is incurred more often.

This interaction of costs has two implications. First, it is possible that a high enforcement effort/low accuracy strategy or a high accuracy/low effort strategy may be appropriate. For many categories of theft, we employ a rather accurate system (which is quite costly per case, given the stakes) and use a low level of enforcement effort (police seriously attempt to capture only a small fraction of thieves).

Second, as will be relevant to the discussion in the next subsection and in Part V, accuracy may be valued for reasons in addition to its effect with regard to controlling behavior. Then, a higher level of accuracy will be appropriate. Given that this is the case, enforcement effort will be more costly at the margin, so the appropriate level of enforcement will be lower. This, in turn, implies that it will be optimal to rely on sanctions to a greater extent, if this is possible. Recall that raising sanctions allows one to reduce enforcement effort in achieving a given level of deterrence. When we wish accuracy to be high, this reduction in enforcement effort is even more valuable. (And, as noted above, raising sanctions and reducing

¹²⁶ The second may add more than the first in an adversary system, if each side hires an expert. But if court appointed experts were used as permitted under Fed. R. Evid. 706, it is possible that one expert would be more accurate than two (where one is appointed by each party).

enforcement effort in a manner that keeps deterrence unchanged does not increase the total amount of mistakenly imposed sanctions.)

Another way to see this difference is to compare two extreme systems, the first with a very high level of enforcement effort (high detection rate) and a low sanction, and the second with a very high sanction and low detection rate. In the first system, accuracy would be extremely costly because there would be so many cases to process, so accuracy would tend to be low. In the second, because there are so few cases, accuracy could be made comparatively high at modest cost—much higher at a given total cost than would be possible with the first system. Thus, the more valuable accuracy is deemed to be, the more one should favor a high sanction, low enforcement effort policy.

2. Accuracy and Sanction Costs

Most of the analysis in Subsection 1 assumed that sanctions were costless, as with monetary sanctions applied to risk-neutral individuals. Consider how the analysis is affected by introducing costly sanctions—nonmonetary sanctions, like imprisonment, or monetary sanctions when individuals are risk-averse.

a. *Sanctions and Enforcement Effort.* It was noted that, when sanctions are costless, there is a benefit of raising sanctions while reducing enforcement effort because a given level of deterrence can be achieved at a lower enforcement cost. Whether sanction costs affect the argument depends on the source of sanction costs. With imprisonment, for example, sanctions are no longer free. But if sanctions are raised and enforcement effort reduced to keep behavior unaffected, the total imposition of sanctions would also tend to be unaffected. (For example, a 50 percent probability of imposing a sanction of 10 and a 25 percent probability of imposing a sanction of 20 both involve an expected sanction of 5.) Thus, raising sanctions and reducing enforcement effort in a manner that keeps behavior unchanged may also keep total sanction costs unchanged while still reducing enforcement costs (because fewer police and court officials are needed if only 25 percent of violators rather than 50 percent are apprehended and processed).

When sanction costs arise because risk-averse individuals are subjected to monetary sanctions, raising sanctions and reducing enforcement effort will raise sanction costs. (In the preceding example, the expected sanction is the same, but the variance is greater when the probability is 25 percent and the sanction is 20.) As a result, it may be optimal to

employ lower sanctions and use greater enforcement effort when individuals are risk-averse.¹²⁷

The presence of inaccuracy does not fundamentally alter this argument. When errors are made, the costly sanctions are less often borne by the guilty and are sometimes borne by the innocent. When sanctions are raised, some of this increase will fall on the innocent rather than on the guilty. But, when enforcement effort is reduced, the overall detection rate is lower, so fewer innocent individuals will be subject to the risk of mistakenly imposed sanctions. As a first approximation, these effects are offsetting.¹²⁸

b. *Accuracy and Enforcement Effort.* In Subsection 1, the discussion emphasized that accuracy and enforcement effort (in addition to sanctions) are substitutes in achieving deterrence, and therefore they should be combined in a manner that achieves a given level of deterrence at the least cost. When sanctions are costly, however, this is no longer the case. Rather, a higher level of accuracy (and thus lower enforcement effort) is appropriate.

The benefit of substituting accuracy for enforcement effort can be seen by comparing how each increases deterrence. Increased enforcement effort enhances deterrence by increasing the detection rate. As a result, more guilty individuals and more innocent individuals will be subject to sanctions. Increasing accuracy enhances deterrence by increasing the number of guilty who are sanctioned while *decreasing* the number of innocent who are sanctioned.

Moreover, if one raises accuracy and reduces enforcement effort so that deterrence is kept unchanged, the total imposition of sanctions on *both* the innocent and the guilty falls. For the innocent, this is straightforward: reducing enforcement effort reduces the number who are detected, and raising accuracy decreases the fraction of detected innocent individuals who bear sanctions. For the guilty, there are conflicting effects: reducing enforcement effort reduces the number detected, but raising accuracy increases the fraction of those detected who bear sanctions. But it must be that the former effect dominates the latter. This follows from the assumption that deterrence is unchanged. For deterrence to be constant, it must be that the difference in expected sanctions for those who commit

¹²⁷ This argument first appears in Polinsky & Shavell, Risk Aversion, *supra* note 123.

¹²⁸ For nonmonetary sanctions such as imprisonment, where risk neutrality may be a good first approximation and sanction costs tend to be linear (see note 111 *supra*), the effects offset precisely. With risk aversion, there are numerous additional subtle effects operating in opposing directions. See Kaplow & Shavell, *supra* note 112.

the act and those who do not remains the same. Since the latter falls, it must be that the former falls as well.¹²⁹

Raising accuracy and reducing enforcement effort a corresponding amount thus reduces the total imposition of sanctions, which is desirable when sanctions are socially costly. Therefore, accuracy should be higher and enforcement effort lower than the level that minimizes enforcement costs.¹³⁰ The most obvious implication is that this provides a rationale for providing relatively more accurate adjudication in the criminal context, where sanctions often involve high social costs, than in the civil context, where sanctions in themselves may involve little social cost.¹³¹ (As will be noted in Section B, the benefit of greater accuracy should be distinguished from any benefits associated with the height of the burden of proof.)

3. Accuracy and Precision in Controlling Behavior

This section has addressed the benefits of accuracy with regard to deterrence and reductions in the imposition of socially costly sanctions, but not Part II's concern with precision in controlling behavior. In this regard, it should be observed that the relationship between deterrence and precision in controlling behavior is closer than may appear. In Sub-

¹²⁹ That is, for any given increase in accuracy, the fall in enforcement effort that keeps deterrence unchanged is necessarily large enough to reduce expected sanctions for the guilty. For if they did not decrease, deterrence would in fact rise because of the fall in expected sanctions for the innocent. See Subsection 1 *supra*.

¹³⁰ In familiar terms, if one is trading off the extent of reliance on two components, adding a benefit on one side of the balance will tip the balance further in that direction. More formally, the levels of accuracy and enforcement effort that minimize enforcement costs are those at which each is equally costly at the margin in producing a given increment of deterrence. Therefore, if accuracy is raised slightly and enforcement effort reduced in an amount that keeps deterrence unchanged, there will be no change in enforcement cost (this is literally true for an infinitesimal adjustment and approximately true for a small adjustment). But this adjustment will produce a positive benefit with regard to sanction costs.

¹³¹ In the criminal context, nonmonetary sanctions (notably, imprisonment) are commonly employed. (In fact, some procedures depend explicitly on the sanction. See *Scott v. Illinois*, 440 U.S. 367 (1979) (counsel must be provided only if imprisonment actually occurs).) In civil cases, sanctions are typically monetary. Even if individuals are risk-averse, the typical sanction cost will be less than in the criminal context per unit of the sanction. With imprisonment, the private cost of the sanction, which affects deterrence, is determined by the loss of liberty and other factors that also constitute social costs—if costs to guilty individuals are deemed to be social costs. To this, one must add the substantial costs of running prisons. Thus, the social cost exceeds the private cost. With risk aversion and monetary sanctions, the risk-bearing cost is necessarily less than the private cost of the sanction (which equals the expected transfer plus the risk-bearing cost). Thus, the social cost is less than the private cost. Therefore, for a given private sanction cost, the social cost is higher for imprisonment than for monetary sanctions.

section IIB1, it is explained that Part II's model concerning damages could be applied to the case of categorizing acts. For example, if some acts were harmful, causing harm of H , and others were harmless, one could interpret this as involving a single set of harmful acts, where harm varied (for some, zero; for others, H). In the language of this section, acts causing no harm could be thought of as innocent. Part II addresses whether it is worthwhile to distinguish acts according to their level of harm. In contrast, the analysis here assumes that there is some value in distinguishing acts, and asks how accurate the legal system should be in making such distinctions.¹³²

In addition, one could supplement the scenario explored in this section in a manner that would make accuracy valuable in achieving more precise control of behavior. Suppose, for example, that there are two types of innocent acts. The first corresponds to that described previously, for which there is a risk of sanctions. The second, equally harmless, innocent act does not subject one to the risk of sanctions. (Perhaps the first type of act must be carried out publicly, in proximity to where harmful acts are regularly committed, while the second type is done privately.)

Consider now the prospect of increasing deterrence of the harmful act by raising sanctions or enforcement effort. Either approach would involve increasing the expected sanction on the first type of innocent act, but without affecting the second type of innocent act. This would induce individuals to favor the second type of act over the first, even when their benefits from the first act were greater. This effect (sometimes referred to as "chilling" innocent behavior)¹³³ would be undesirable.

If, instead, deterrence were enhanced by increasing accuracy, the expected sanction for the first type of innocent act would fall. This would reduce the extent to which individuals are inefficiently induced to choose the second act over the first. Thus, achieving more precise control of behavior can be a benefit of accuracy, in addition to enhancing the level of deterrence for harmful acts as a whole.¹³⁴

¹³² The case from Part II in which damages equaled average harm would, in this section, correspond to one in which both the innocent and guilty were equally likely to be apprehended and in which the rate of correct sanctioning of the guilty just equaled the rate of false convictions of the innocent. Part II implicitly assumed that the only alternatives were complete accuracy in making the distinction and not making any distinction. In this section, imperfect distinction is allowed. One can control behavior just as well in such cases as long as expected sanctions are high enough. But if there are limits to or costs of raising sanctions and enforcement effort, expenditures to enhance accuracy are desirable.

¹³³ See, for example, Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the "Chilling Effect,"* 58 B.U.L. Rev. 685, 694-701 (1978).

¹³⁴ It may be possible, however, to achieve this benefit in an alternative manner. Png notes that the first type of innocent act could be subsidized. See Png, *supra* note 118. This

B. Burden of Proof

1. Relationship between Burden of Proof and Accuracy

The previous analysis assumed that designing the legal system to regulate a type of activity involved choosing the level of the sanction, enforcement effort, and accuracy. This formulation implicitly takes the burden of proof as given. But the proof burden can be chosen as well. This need not, however, affect the preceding discussion, which indicates how to choose the sanction, enforcement effort, and accuracy for any burden of proof that may be specified.¹³⁵ Hence, determining the burden of proof could be left as a separate inquiry. It is the case, however, that setting the burden of proof and choosing the level of accuracy often are related.

Initially, it is useful to define a shift in the burden of proof in isolation—that is, taking as given all other aspects of the system, including its accuracy. When a decision-maker hears all the evidence, there inevitably remains some uncertainty about what actually happened. A higher proof burden is taken here to mean that the decision-maker must have a higher level of confidence that the party is responsible for committing the illegal act in order to impose a sanction. Thus, raising the burden of proof increases the rate of mistaken acquittals and decreases the rate of mistaken convictions.

In contrast, increasing accuracy was said to reduce the rate of mistaken acquittals and convictions. This arises when better information is available to the decision-maker or a more capable decision-maker is used. For a given burden of proof, better information or better analysis of given information reduces mistakes of both types. (Assume that the burden of proof is held constant by setting the required level of confidence at a point such that the same percentage of convictions and acquittals results. Then the rates of both types of error necessarily fall or rise together as information improves or worsens.)

The relationship can be explored further with an example. Consider a policy of subsidizing defendants' legal counsel. Assume for present pur-

might be accomplished either by providing some reward to those found innocent (in an amount that, *ex ante*, just offset the possibility of being mistakenly sanctioned) or by subsidizing the activity directly.

¹³⁵ It has been argued that the questions may not be independent, in that juries may choose to require more proof if sanctions are higher. See, for example, James Andreoni, *Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?* 22 *RAND J. Econ.* 385 (1991) (offering a model in which the cost of false convictions rises with the sanction but the benefit of correct convictions—as by increasing deterrence—is stipulated not to increase with the sanction).

poses that the policy increases the accuracy of the proceedings.¹³⁶ Typically, it is assumed that such a policy also enhances both innocent and guilty defendants' chances of success. In the terminology employed here, this policy involves both an increase in accuracy and an (implicit) increase in the burden of proof. (Thus, the burden of proof is interpreted *de facto* rather than with regard to its legal formulation: if the probabilities of conviction for the innocent and the guilty are lower for whatever reason, the proof burden is said to be higher.)

Consider some alternative policies. Suppose there is a subsidy of defendants' legal counsel and a simultaneous downward shift in the express burden of proof to an extent that the portion of detected individuals who are ultimately convicted remains the same.¹³⁷ Then, the only effect would be on accuracy.

The question thus becomes whether the additional accuracy is worth the cost, the burden-of-proof question being disposed of separately. For example, what if increasing resources available to the prosecution by the same amount increased accuracy more? Then, one could accompany this change by an increase in the express burden of proof, producing more accuracy at the same cost. Similarly, if what one really desired in advocating a subsidy to defendants' legal counsel was making the *de facto* burden of proof higher, one might accomplish this directly or through other indirect means (such as by reducing resources available to the prosecution). Thus, it is possible in principle to view accuracy and the burden of proof separately.¹³⁸ Any policies affecting both can be analyzed with regard to each component because if one aspect is desirable and the other

¹³⁶ This is obviously a controversial assumption. The purpose of the discussion in the text is to clarify what is meant by increased accuracy and to describe how one should analyze the policy, whatever one thinks to be the truth about its effect on accuracy.

¹³⁷ In reality, there would be no need for an explicit shift if the decision-maker discounted the information presented to account for the presence of defense counsel (or, equivalently for present purposes, discounted the prosecution's case when there is no defense counsel). A sophisticated decision-maker who understands how the process works would be aware of the effect of counsel on the average tilt of the information presented and thus could not help but consider separately the question of the burden of proof.

¹³⁸ Some rules may be related to accuracy and the burden of proof in a special manner. Namely, constitutional protections (such as the exclusionary rules of the Fourth and Fifth Amendments) may be designed to prevent abuse of power by the government. Many applications of these protections may decrease accuracy, and even if accuracy is increased (as by prohibiting the admission of confessions obtained under questionable, but not entirely unreliable, circumstances), there may be other ways to increase accuracy at a lower cost. Yet the provisions may exist to make it difficult for the government to go after political opponents or unpopular individuals. It may be feared, for example, that high proof burdens and generally adequate procedures would sometimes be insufficient. See Subsection VC3 and note 271 *infra*.

not, it would be possible to benefit from the former and do without the latter.¹³⁹

2. Setting the Burden of Proof¹⁴⁰

a. *Controlling Behavior.* The burden of proof affects the control of behavior. Obviously, an infinite burden or a zero burden would tend to be disastrous, as adjudication then would not differentiate according to behavior.¹⁴¹ Ignoring any effect on sanction costs, the optimal burden of proof would be that which maximizes deterrence. The reason is that adjusting the burden of proof is free,¹⁴² while enforcement effort and

¹³⁹ What is true in principle may sometimes be impossible in practice. In the civil context, there is probably substantial room to modify the express burden of proof directly, and this is sometimes done. In the criminal context, one might argue that it is impossible to increase further the express burden of proof (and express decreases may be constitutionally impermissible) because we already have the highest possible express burden (short of requiring absolute certainty, which would require acquittal in every case). This appearance, however, is probably misleading. First, indirect means will fail in any event if decision-makers are sophisticated. See note 137 *supra*. Second, I am unaware of any demonstration that it is not possible to adjust the burden upward. (After all, the current formulation seems nearly absolute, yet conviction is routine on evidence falling well short of certainty; this may be in part because sophisticated fact-finders do discount for aspects of the process or simply because fact-finders disregard instructions. One can imagine empirical research that would illuminate this question further.) Third, there are many ways of raising proof burdens. For example, one can adjust or add to elements of offenses. (One can heighten the required intent or make aspects of proof elements of a crime—for example, the two-witness requirement for treason. See also Note, *Winship on Rough Waters: The Erosion of the Reasonable Doubt Standard*, 106 Harv. L. Rev. 1093 (1993).) For a possible exception, see note 138 *supra* (rules that implicitly raise the proof burden to avoid abuse of government power). Finally, note that changes in the burden of proof affect litigants' incentives to present information and thus affect accuracy.

¹⁴⁰ Prior economic analyses that have considered the burden of proof include Johnston, *supra* note 43; Posner, *supra* note 118, at 408–15; Daniel L. Rubinfeld & David E. M. Sappington, *Efficient Awards and Standards of Proof in Judicial Proceedings*, 18 RAND J. Econ. 308 (1987) (studying how the burden of proof affects defendants' litigation expenditures). See also Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions* (unpublished manuscript, Harvard Univ., Kennedy School of Gov't, 1992) (discussing proof burdens outside the formal legal setting). Earlier advocates of a cost-benefit approach to determining burdens of proof include Alan D. Cullison, *Probability Analysis of Judicial Fact-finding: A Preliminary Outline of the Subjective Approach*, 1969 U. Tol. L. Rev. 538, and John Kaplan, *Decision Theory and the Factfinding Process*, 20 Stan. L. Rev. 1065 (1968). See also Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 Harv. L. Rev. 1329, 1378–93 (1971) (critiquing such an approach, although not advocating a concrete alternative).

¹⁴¹ A zero burden does not really exist in most contexts even if there is a zero burden in adjudication. For example, if the police apply a threshold in apprehension (rather than making entirely random arrests), the *de facto* burden is not really zero. This illustrates how accuracy involves the combined effect of choices in investigation (or auditing, monitoring, etc.) and in adjudication. See note 116 *supra*.

¹⁴² For qualifications, see note 139.

accuracy are costly. Therefore, if it is possible to raise deterrence by altering the burden of proof, one could make such a change in the proof burden while simultaneously reducing, say, enforcement effort. This would keep behavior unchanged while reducing enforcement costs.¹⁴³

What burden of proof would maximize deterrence? Not one that was very low. To see this, note that reducing the burden of proof not only raises the likelihood of sanctioning the guilty but also that of sanctioning the innocent. When the burden is low, most of the guilty (for whom there is any significant evidence of guilt) may be found guilty in any event, in which case a further reduction would primarily increase the rate at which the innocent are convicted. As explained in Subsection A1, this would reduce deterrence. Just what burden of proof results in maximum deterrence is difficult to determine. It might be thought that a "more likely than not" standard, common in the civil context, maximizes deterrence. The idea is that such a threshold is precisely the one below which it is more likely that one is sanctioning innocent behavior. But the relationship between the proof burden and behavior is more complicated.¹⁴⁴

b. *Sanction Costs.* When sanctions are themselves costly, as with imprisonment, a common intuition is that a higher burden of proof is appropriate. This subsection offers an argument suggesting that this intuition is correct. Increasing the burden of proof reduces sanctions on both the innocent and the guilty and therefore reduces sanction costs. This might seem to be enough of a demonstration: the greater the social cost of sanctions, the greater the benefit of increasing the burden of proof.

There is, however, an important complication. Increasing the burden

¹⁴³ Compare Subsection A1a *supra* (optimal to raise sanctions to maximum feasible extent if sanctions are costless).

¹⁴⁴ To illustrate, a lower burden could readily improve behavior. Suppose, for example, that when a harmful act is committed, the injurer cannot be distinguished from the two individuals in closest proximity. Then, the probability that one identified the correct injurer would always be one-third. A proof burden of "more likely than not" would exonerate everyone. A lower burden would sanction all injurers as well as additional individuals. But this may have little adverse effect on behavior. First, if the sanction is sufficiently high, no one may commit the act in such circumstances. Second, as long as the acts are not terribly frequent, innocent individuals would not have their behavior much affected by the slight possibility that they would be the unlucky parties who happen to be those nearest to the injurer.

Alternatively, if an act is committed by many and there is some difficulty in identifying the injurers accurately, it may be that many innocent individuals would find themselves in situations—proximate to others committing the act—where it would seem in adjudication more likely than not that they, too committed the act. Then, they would face no increased risk of sanction by committing the harmful act, an undesirable state of affairs. But it may be that requiring a higher standard of proof would usually exonerate such individuals, assuming they did not act, while still resulting in a sanction being applied to most who actually committed the act.

of proof also affects deterrence. Suppose that, in the range under consideration, an increase in the proof burden reduces deterrence.¹⁴⁵ Then more individuals commit the harmful act. But individuals who commit the harmful act are more likely to be sanctioned than individuals who do not.¹⁴⁶ This effect increases the total imposition of sanctions. In total, a lower fraction of the innocent and guilty are sanctioned, but a higher fraction of the population will be guilty and thus be sanctioned more often.

Although the effect on deterrence complicates the argument for a higher burden of proof, the argument is not undermined entirely. The analysis in Subsection a indicated that, when sanctions are costless, the most efficient burden of proof is that which maximizes deterrence. If one increased the proof burden slightly above that point, the rate at which deterrence would fall is generally very small.¹⁴⁷ In contrast, raising the burden would reduce sanction costs at a positive rate.¹⁴⁸ Therefore, raising the burden of proof somewhat above the maximum deterrence point will tend to reduce sanctions, and therefore sanction costs, without significantly affecting deterrence. One could continue to raise the burden further, but at some point further increases will reduce deterrence more significantly. That reduction will be undesirable both because it offsets to some extent the direct effect of reducing the rate of imposition of sanctions and because behavior is less effectively controlled.¹⁴⁹ The most

¹⁴⁵ If the increase in burden of proof increases deterrence, it is probably desirable because there is both a deterrence benefit (but see the qualification in note 149) and a reduction in sanction costs. One would stop increasing the burden further only at some point after deterrence began to fall.

¹⁴⁶ Otherwise, there would be no deterrence whatsoever.

¹⁴⁷ The logic is analogous to that in note 130 *supra*. Intuitively, the argument is that, in most policy settings, little is lost as a result of being near the ideal policy rather than precisely at it. Thus, increasing the proof burden from the level that maximizes deterrence initially has a negligible effect on deterrence but not on the imposition of sanctions. Formally, the argument is one from calculus: a function has a zero derivative at its maximum, meaning that the rate of change (here, in deterrence) is zero at the maximum and close to zero when near the maximum.

¹⁴⁸ The emphasis is on rates, rather than amounts. This relates to the point in note 147.

¹⁴⁹ This latter effect depends on the assumption that, at the optimum—taking into account sanction costs, enforcement costs, and control of behavior—there tends to be underdeterrence. This is true when one considers only enforcement costs and behavior. (The reason is that the benefit of deterrence with regard to behavior alone is subject to diminishing returns, and zero marginal returns precisely at the ideal level of deterrence, while the cost of marginal increases in deterrence remains positive.) When sanctions involve social costs, however, it is possible in principle for optimal sanctions to involve overdeterrence. See Louis Kaplow, A Note on the Optimal Use of Nonmonetary Sanctions, 42 J. Pub. Econ. 245 (1990); Louis Kaplow, The Optimal Probability and Magnitude of Fines for Acts That Definitely Are Undesirable, 12 Int'l Rev. L. & Econ. 3 (1992); Polinsky & Shavell, *supra* note 111.

efficient burden of proof will be that which just balances these opposing factors.¹⁵⁰

It is possible to say something more concrete about the ideal burden of proof when sanctions are socially costly. The preceding argument suggests that, when the burden of proof is set properly, it will be at a level at which deterrence is falling. Thus, at the margin, a rise in the burden of proof (say, the last increment before reaching the optimal level) is more generous to the guilty than to the innocent. This is consistent with the view that the criminal justice system should err on the side of being generous to the guilty for the sake of avoiding the imposition of sanctions on the innocent, although the correspondence is hardly precise.¹⁵¹

Finally, note how the analysis of the burden of proof combines with that of accuracy in the case of costly sanctions. Subsection A2b indicates that more costly sanctions warrant a greater level of accuracy, which in turn may best be accomplished with a lower level of enforcement effort and higher sanctions (because greater accuracy raises the cost of enforcement effort). This subsection suggests that, when sanctions are more

¹⁵⁰ The analysis in the text shows that with costly sanctions the burden of proof should be higher than with costless sanctions. A corollary is that, the greater the sanction cost, the higher the proof burden should be. (For reasons discussed previously, at the most efficient proof burden for a given sanction cost, one would expect that a marginal increase in the proof burden would reduce deterrence. For that proof burden to be optimal, it must be that there is a reduction in total sanction costs just sufficient to offset the adverse effect with respect to controlling behavior. Then, if the social cost per unit of the sanction were higher, a marginal increase in the burden of proof would be desirable, rather than simply equating marginal costs and benefits.)

¹⁵¹ The correspondence is imperfect for two reasons. First, the analysis in the text, without empirical evidence, does not indicate the extent to which the guilty should be favored (while it is commonly suggested that the extent is quite substantial). Second, the generosity toward the guilty suggested by the argument in the text is in terms of expected sanctions. When deterrence falls, it is because the fall in expected sanctions for the guilty exceeds the fall in expected sanctions for the innocent. Most commentators, however, probably have in mind not the expected sanctions for the innocent, but the treatment of the innocent who are arrested and subject to the formal legal process. Observe that a shift in the proof burden that permits, say, one additional guilty person and one additional innocent person go free is favorable to the guilty on an expected value basis, as long as less than half of the relevant population commits the act. (Because there are fewer who commit the act, the expected sanction each guilty person faces must fall by more to produce an expected reduction in the number of guilty sanctioned of one, which equals the reduction for the innocent, of whom there are more.) Thus, the logic of the argument in the text, suggesting that generosity toward the guilty in terms of expected sanctions is efficient, does not imply that the proof burden necessarily must be high enough that, at the margin, more guilty go free than innocent.

The lack of a precise fit may not be surprising if the burden of proof in criminal cases is justified in part by noneconomic concerns, such as those explored in Section VB. (To foreshadow that analysis a bit, note that the "economic" concerns here include the social cost of depriving individuals, both innocent and guilty, of their liberty.)

costly, proof burdens should be higher, to an extent that reduces deterrence. It may be appropriate to address some of the resulting shortfall in deterrence by other means. In light of the comments on accuracy, sanctions, and enforcement effort, it may be best to rely most on enhanced accuracy and higher sanctions.¹⁵² Viewed together, when sanction costs are high, it seems that high proof burdens, high accuracy, high sanctions, and a low rate of enforcement may be an appropriate mix. This combination characterizes criminal law to a substantial extent, in contrast to the civil context.¹⁵³

C. Applications

1. Civil Cases and Private Dispute Resolution

As suggested at the outset of this part, the analysis is applicable to all dispute resolution contexts even though the language speaks of the innocent and guilty and the use of sanctions. While some examples outside the criminal context were noted in some instances, it is useful to consider further how the analysis applies.

Accuracy in determining liability and the burden of proof are features of all legal settings involving the control of behavior. In place of criminal sanctions, damage awards are more common in civil and private contexts, but the logic is unaffected. Corresponding to the level of criminal enforcement effort are aspects of the civil legal system that affect incentives to sue. With crimes, where the government has a legal monopoly of enforcement, it is possible to bring cases that, *ex post*, may not appear cost-justified (for example, because the costs of trial or prison exceed the harm) in order to enhance deterrence. In contrast, private litigants typically have no such incentive.¹⁵⁴ Moreover, the government can refrain

¹⁵² Because this strategy decreases the rate of imposing sanctions, in contrast to using more enforcement effort, it lessens the need to raise the proof burden; similarly, a higher proof burden, by reducing the imposition of costly sanctions, makes accuracy less valuable at the margin.

¹⁵³ It is familiar that in civil proceedings sanctions are less socially costly (see note 131 *supra*), less effort is made to be accurate, proof burdens are lower, and sanctions are often lower (depending on how one compares damage awards and imprisonment). With regard to enforcement effort, note that in civil cases the injurer can often be identified and the victim often has an incentive to sue, as long as the stakes are not low relative to litigation costs. Thus, the probability of suit in many settings is high. For the criminal law, even when harm is substantial, this is not true (aside from some exceptional areas, such as murder).

¹⁵⁴ An exception arises when they contemplate future interactions and can benefit from a reputation of suing even when the stakes are small, in order to deter violations or to induce settlements before significant litigation costs are incurred.

from pursuing highly meritorious cases if it has decided that additional prosecutions would not make a sufficient contribution to deterrence.¹⁵⁵ In contrast, private litigants will pursue cases when the expected recovery exceeds the cost of litigation, even if there would be no contribution to deterrence.¹⁵⁶

While the government does not control each litigation decision in civil suits, it does structure the process in a manner that affects incentives to sue. Litigation costs may be subsidized or shifted. Alternatively, fees may be charged and penalties imposed (as under Rule 11). Also, in the private context, there is a direct interaction between sanctions (damages) and incentives to sue: raising sanctions tends to encourage litigation. Although adjusting sanctions is less common in the civil context (in most areas of law, damages equal harm), there are exceptions (treble damages, punitive damages, rules limiting consequential damages, requirements concerning proof of damages), and more could be made.¹⁵⁷ Also, public enforcement in the civil context is important, whether exclusive or supplemented by private rights of action.¹⁵⁸ Similarly, with private dispute resolution (or civil litigation enforcing contracts), parties can adjust the system in various ways, most obviously by specifying liquidated damages (which affects litigation costs, incentives to sue, and the deterrent effect of a given suit).

The major difference between public enforcement in the criminal context and enforcement in other contexts thus involves the nature of sanctions and the typical degree of directness in controlling the frequency of enforcement.¹⁵⁹ Even here, the difference could be reduced. Some commentators have proposed schemes such as decoupling the damages paid by defendants from awards collected by plaintiffs.¹⁶⁰

¹⁵⁵ Of course, the government may also consider other benefits, such as incapacitation (addressed in Subsection IVB2).

¹⁵⁶ The discussion in Section IIC involved this sort of problem. See A. Mitchell Polinsky, *Private versus Public Enforcement of Fines*, 9 J. Legal Stud. 105 (1980); Shavell, *supra* note 94.

¹⁵⁷ Polinsky and Shavell, *supra* note 118, consider how setting damage awards and payments by losing plaintiffs affects incentives to sue in a model in which there is error in determining liability.

¹⁵⁸ For example, enforcement of much modern regulation involves public prosecution in a civil setting.

¹⁵⁹ As emphasized in Section IIC, accuracy is chosen by private parties—in the civil context and in the criminal context (with respect to defendants). Thus, accuracy is determined indirectly, through procedural rules, rules of evidence, proof burdens, and definitions of substantive law.

¹⁶⁰ See A. Mitchell Polinsky & Yeon-Koo Che, *Decoupling Liability: Optimal Incentives for Care and Litigation*, 22 RAND J. Econ. 562 (1991). See also A. Mitchell Polinsky &

2. Liability versus Regulation

In order to control behavior, the legal system varies its reliance on liability (as with tort) and regulation (as with traffic or pollution laws), sometimes relying almost entirely on one or the other and at other times using both. Although the many factors relevant to the use of liability and regulation have been examined elsewhere,¹⁶¹ it is useful here to note how the present analysis bears on their use.

Regulation is usually directed at behavior before harm arises. Regulation may be applied to each actor's behavior (as when all restaurants are inspected), or it may rely on probabilistic enforcement (as with traffic violations). Sanctions in many contexts are rather modest, reflecting in part that violations may not cause harm but only indicate a small probability that harm might have been caused.¹⁶² With liability, in contrast, it is typical to wait for harm to occur, allowing the victim to sue to recover damages equal to the amount of harm caused. In accident cases, the ex ante probability that harm will occur is often low, and damage awards are high in comparison to fines for equivalent regulatory violations. (For example, the fine for speeding may be \$50, but if one who is speeding hits another car, liability may range from thousands to millions of dollars, depending on how much damage results.)

The prior analysis might be taken to suggest that liability is superior for two reasons. First, by using a high sanction with a low probability,

Daniel A. Rubinfeld, Optimal Awards and Penalties When Some Suits Are Frivolous (Working Paper No. 93, Stanford Law School, John M. Olin Program in Law and Economics 1992) (increasing damage awards while penalizing losing plaintiffs may reduce the rate of litigation while maintaining deterrence). Such alternatives may best be viewed as extending Becker's intuition, see Becker, *supra* note 122; Subsection A1a *supra*, to the civil context involving private enforcement. See Louis Kaplow, Shifting Plaintiffs' Fees versus Increasing Damage Awards, 24 RAND J. Econ. 625 (1993).

¹⁶¹ See Shavell, *supra* note 74; Steven Shavell, The Optimal Structure of Law Enforcement, 36 J. Law & Econ. 255 (April 1993).

¹⁶² There still remains the question of why fines are not higher, with a lower probability of detection. See Subsection A1a *supra*. For example, fines for common traffic violations may be about \$50, and one could imagine fines being raised to \$100, allowing enforcement effort to be cut in half. The probability may be kept as high as it is (and thus fines may be low) for many reasons. First, the police who monitor traffic violations are also in a position to respond to emergencies and to watch for other illegal acts. See Shavell, *supra* note 123. (Still, substantial police resources are devoted solely to traffic violations, as when many officers staff a speed trap.) Second, if the probability were too low, misperceptions about its magnitude might become relatively more significant. See Bebachuk & Kaplow, *supra* note 123. Third, if there are too few monitors, it will often be the case that potential violators will be certain (by direct observation) that no police are nearby, so that they can violate the law with impunity. (Yet one suspects that this is often the case in any event, and changing the number of traffic police would not greatly affect how often this is true.)

enforcement resources are saved.¹⁶³ Second, it might appear that tort liability is superior because of accuracy. When the probability is so much lower, one can invest more in each case to promote accuracy. Although this is true all things being equal, the differences between liability and regulation suggest that the latter may in fact be more accurate. When the police issue speeding tickets, they either measure the speed with radar or directly observe the car (and they are experienced at making such observations, focusing on the relevant question, and making contemporaneous notes of what they see when they write a ticket). In contrast, in an auto accident case, witnesses who are less expert, have no radar equipment, may not have been paying attention, and often are interested parties offer later testimony about the speed of a car.¹⁶⁴ Because differences inherent in the way information is gathered have such a large effect on accuracy, it is unlikely that subsequent expenditures could eliminate the disparity in accuracy in these contexts.

D. Individuals' Information at the Time They Act and Their Incentives to Present Information in Adjudication

1. Individuals' Knowledge Ex Ante and the Value of Accuracy Ex Post

Much of the analysis in Part II emphasizes individuals' knowledge of legal consequences at the time they decide how to act because accuracy ex post is valuable only if individuals will be induced to behave in a manner that reflects the distinctions that will be made in adjudication. Thus, if an individual is unaware whether her act will cause more or less

¹⁶³ One might save further resources by, for example, raising damage awards in auto accident cases while simply eliminating police monitoring. The judgment-proof problem is the most important of many reasons for not doing so. See also note 162 *supra*. Alternatively, one might eliminate tort recoveries for auto accidents and correspondingly increase the level of traffic fines, maintaining deterrence while saving the costs of the tort system (using first-party insurance for compensation). This may have two problems. First, when one is confident that police are not monitoring an area, there is no deterrence. Second, because tort awards depend on harm caused, they will reflect the degree to which behavior is dangerous. (But traffic fines could be more finely graded, such as by raising the fine for driving over the center line, having fines vary more precisely as a function of speed and traffic density, and so on. Moreover, as emphasized in Part II, much of the precision with regard to ex post damage determinations is not anticipated ex ante. Thus, one running a red light at forty miles per hour will have little idea of who if anyone will be hit, how badly potential victims would be injured, and the like. A fine based on statistical data for such acts could provide equivalent incentives.)

¹⁶⁴ The testimony may be years later, but often there will have been interviews soon after the accident. Even so, an interview days or even hours later does not compare to contemporaneous responses of a careful observer concentrating on the event in question.

harm than average, damages based on actual harm rather than average harm will not affect her decision how to act.

Similar reasoning is applicable to accuracy in determining liability. Thus, if an individual does not know which of two acts she contemplates committing is harmful, greater accuracy in distinguishing the acts *ex post* will not affect her behavior. To illustrate, suppose that, with perfect accuracy, the expected sanction for the first act would be 100 and, for the second, 0. If one does not know which act is subject to the sanction of 100 and which 0, the expected sanction for each would be 50.¹⁶⁵ Alternatively, if there were error, so that the expected sanction for the first act was 80 (because 20 percent of the guilty are exonerated) and the expected sanction for the second was 20 (because some who commit the innocent act are mistakenly sanctioned), the expected sanction for each act would still be 50. Thus, only if the actor knows which act is harmful, and thus legally subject to sanctions, will accuracy *ex post* be relevant for behavior.¹⁶⁶ If actors initially are uninformed, greater accuracy *ex post* will increase the incentive to acquire information *ex ante*, before deciding how to act.

Although it is entirely natural to assume in the context in Part II involving damages that individuals often would not know *ex ante* how much harm their act would cause *ex post*, it seems more often to be the case with regard to liability that individuals know whether their acts are illegal. This seems most likely with regard to common crimes. For some violations of modern statutes (tax laws, occupational safety requirements), however, it seems plausible that many actors would have substantial uncertainty about which acts are deemed harmful by the legal system. The more it is the case that individuals are uninformed *ex ante*, the less

¹⁶⁵ The example uses the assumption that the actor believes that there is a 50 percent chance that each act is subject to the sanction of 100. If individuals believed that the chance that the first act was harmful exceeded 50 percent, there would be some deterrent effect of accuracy, although less than when individuals are certain. (If individuals believed that the truly harmful act was the harmful one with a probability under 50 percent, then more accuracy would actually make behavior worse.)

¹⁶⁶ This is literally true only if the assumption is that all are uninformed. Assume that some individuals are informed and some uninformed, that the adjudicator does not know who is who, and that there is error. It may be that it would be desirable to raise expected sanctions on account of the informed because expected sanctions of 80 versus 20 involve a dilution of deterrence, so raising sanctions or enforcement, perhaps producing expected sanctions of 130 and 30, might seem desirable. But then expected sanctions would distort behavior if there were additional options. See Subsection A3 *supra*. For an analysis of uncertainty with regard to liability in a model in which error is not involved, see Louis Kaplow, Optimal Deterrence, Uninformed Individuals, and Acquiring Information about Whether Acts Are Subject to Sanctions, 6 J. L. Econ. & Org. 93 (1990).

valuable accuracy will be *ex post*.¹⁶⁷ Given that the discussion of liability in Part III can be seen as an instance of uncertainty about damages in Part II,¹⁶⁸ it should not be surprising that the insights of Part II would be applicable in this part.

2. Whether Individuals' Incentives to Present Information in Adjudication Are Socially Appropriate

a. *Costless Sanctions*. Section IIC emphasized the divergence between parties' incentives to present information about damages in adjudication and the social value of information in that setting. A similar question can be asked with regard to information about liability. As the preceding subsection suggests, the same analysis will be applicable. To the extent individuals are uninformed *ex ante* about whether their acts are subject to liability, greater accuracy *ex post* will have no social value, but presenting information will be valuable to parties to the extent of its effect on the expected outcome of litigation. And if individuals are informed *ex ante*, incentives *ex post* may still be inappropriate because the effect of accuracy *ex post* on behavior *ex ante* does not tend to equal the private benefit to parties *ex post* in presenting information in adjudication.

b. *Costly Sanctions*. The analysis in Section IIC of parties' incentives to present information in adjudication viewed the effect of such information on the expected outcome as being of value only indirectly, through its effect on behavior. When sanctions are socially costly, changes in the outcome also involve direct costs. The present discussion will consider expenditures by innocent individuals made to convince a tribunal of their innocence.¹⁶⁹

When innocent individuals provide better information, decreasing the

¹⁶⁷ When some individuals are informed *ex ante* and others are not, it is natural to consider whether different sanctions should be applied to these two groups. See *id.* (suggesting that there is no reason for differentiating the two groups unless there are systematic misperceptions by the uninformed or sanctions are socially costly). If the distinction is to be made, the related question of how accurate adjudication should be in making the distinction is also presented.

¹⁶⁸ See Subsection A3 *supra*.

¹⁶⁹ As in Section IIC, attention here is confined to the presentation of information that improves the accuracy of the result. The truly guilty can only improve the outcome by convincing the tribunal of their guilt, which they have no incentive to do. In contrast, in Part II, the truly guilty may have caused less harm than the court would think were they to remain silent. But, as Subsection A3 indicated, there are analogues in the present context. For example, the truly guilty may be guilty of a lesser offense than the tribunal would otherwise think. But, using what the tribunal initially believes as a benchmark, such individuals should be viewed as "innocent" for present purposes.

chance of a mistake, there is both an improvement in deterrence (the anticipation that mistaken sanctions will be imposed on the innocent less often makes innocent acts relatively more attractive) and a reduction in sanction costs. Although the analysis in Section IIC suggests that the significance of the former factor alone (the degree of improvement in *ex ante* behavior) may be less than the change in the expected sanction (which determines an individual's incentive to present the information), the balance would be otherwise if the social costs of sanctions were sufficiently large. When the sanction is imprisonment, moreover, it seems that social sanction costs alone exceed private sanction costs. The reason is that the private sanction cost—loss of liberty—is a social cost, whereas the substantial cost of running prisons is a social cost but not a private cost. Thus, the innocent would have too little incentive to prove their innocence, as surprising as this may seem. (And this is true even ignoring the fact that reducing mistaken convictions of the innocent has the further social benefit of improving behavior.)

When sanctions involve high social costs, as they often do in the criminal context, economic considerations thus may favor subsidizing defense expenditures of the innocent.¹⁷⁰ There is, however, a problem. The very reason for criminal trials is that in advance we are not certain who is innocent. If, instead, one subsidizes all criminal defendants, the guilty—who already have an excessive incentive to prove their innocence¹⁷¹—are further encouraged. One would also be concerned about whether accuracy overall would increase or decrease as a result; if the latter occurred, the final result may involve more mistaken convictions of the innocent.¹⁷²

One can imagine other ways to attempt to address the problem.¹⁷³ For example, if criminal defendants who were exonerated received rewards, there would be more of an incentive to undertake litigation expenditures, but the increase in incentive may be relatively greater for the innocent. Alternatively, one could subsidize criminal defense but penalize those

¹⁷⁰ See also Subsection B1 (discussing the example of subsidizing defense counsel for all defendants if such a subsidy would increase accuracy).

¹⁷¹ See note 169 *supra*.

¹⁷² As discussed in Subsection B1, subsidizing defense counsel would shift the burden of proof and affect accuracy. The two components can be separated. Thus, if accuracy falls, then for any given level of mistaken acquittals of the guilty, there will be more mistaken convictions of the innocent. See also note 137 *supra* (indicating that, even if the formal burden of proof is not adjusted, sophisticated decision-makers may implicitly adjust it in any event). If one valued defense counsel, without regard to the effect on accuracy, solely to reduce mistaken convictions of the innocent, it would be better simply to raise the burden of proof. (This would not be true if defense counsel were of more value to the innocent, but this is precisely the case in which accuracy would improve.)

¹⁷³ The mentioned possibilities are hardly exhaustive. Thus, one might imagine providing an added subsidy to the defense of those with no prior record or those who pass a polygraph test—factors that may be probative of guilt although inadmissible at trial.

ultimately convicted, and presumably the guilty have a greater likelihood of conviction. The present criminal justice system has this latter feature, for the indigent (a large portion of criminal defendants) receive subsidized services, but if one uses them to a great extent by going to trial,¹⁷⁴ it is generally believed that one's sentence is likely to be greater than if one pleads guilty beforehand.¹⁷⁵

IV. ACCURACY IN ESTABLISHING FUTURE RIGHTS AND OBLIGATIONS

The analysis in Parts II and III considers settings in which the primary purpose of adjudication is to provide incentives to guide prior conduct.¹⁷⁶ In such contexts, sanctions themselves are of no independent interest except to the extent they are socially costly. In that case, it is desirable to design a legal system in a manner that reduces the use of sanctions, and an important value of accuracy in adjudication is that it has such an effect.

In this part, the analysis will focus on situations in which the outcome of adjudication is not merely a sanction, or not a sanction at all, but rather a pronouncement of future rights and obligations that have a direct effect on the future actions of the parties rather than an indirect effect through deterrence. Thus, it will be of independent significance whether adjudication is accurate *ex post*, and the notion that sanctions should be kept to a minimum will no longer make sense.

Section *A* considers adjudication of future entitlements, an important instance in which the primary concern of adjudication is prospective. Section *B* addresses adjudication about past acts that governs future conduct, in which the analysis in Section *A* as well as that in prior parts of this article is relevant.

A. Adjudication about Future Entitlements

Many disputes are primarily, if not exclusively, concerned with future entitlements. For example, licensing procedures may determine whether one is fit to practice a profession or perform a service.¹⁷⁷ Another example, which will be the focus of this section because of the attention it has

¹⁷⁴ Guilty pleas obviously save prosecutorial and court resources in addition to defense resources.

¹⁷⁵ See, for example, Kenneth Jost, *A Six-Month Burden on Right to Trial?* A.B.A. J. 34 (June 1993). One also suspects that the increase in sentence for going to trial tends to be greater when it is apparent at the time of sentencing that the defendant was obviously guilty and knew it, rather than involving a case with an arguable defense (such as lack of intent for tax evasion or possible self-defense).

¹⁷⁶ One exception is the discussion in Subsection II.A.3 that considers whether damages provide accurate compensation to risk-averse plaintiffs.

¹⁷⁷ Licensing may also involve past behavior. In license renewal proceedings, this is

received,¹⁷⁸ involves the entitlement to future public payments,¹⁷⁹ as from welfare programs.¹⁸⁰ Subsection 1 begins the analysis by offering an assessment of the value of accuracy in this context. Subsection 2 addresses whether claimants' incentives to make expenditures in adjudication to establish their entitlements are socially excessive. Subsection 3 discusses how the present analysis differs from that most commonly offered by courts and commentators.

1. The Value of Accuracy¹⁸¹

Consider a scenario in which individuals apply for disability benefits.¹⁸² Eligibility for benefits depends on whether one is disabled.¹⁸³ The system

obvious. In addition, if licenses are denied to applicants with, say, prior criminal convictions, the license denial can be seen as part of the sanction for prior acts.

¹⁷⁸ What is surprising is that most of the other contexts addressed in this article have received so little attention.

¹⁷⁹ Past behavior may be involved here as well. For example, if one quit one's job rather than being fired, one may be ineligible for unemployment compensation. Moreover, many benefits affect ex ante incentives—to earn income in the case of welfare programs or to avoid injury in the case of social security disability insurance. These effects are ignored for convenience. (The latter effects have the most general application. The decision to have the program reflects a judgment that payment of benefits to individuals who qualify is desirable in spite of these incentive effects. To incorporate the incentive effects, one may simply deem subsequent references to the benefits of providing payments to qualified individuals to be the net benefits: the benefit of compensation minus the incentive cost.)

¹⁸⁰ Other contexts involving constitutional claims concerning due process—such as school discipline, being fired from government employment, or parole revocation—raise analytically similar issues.

An important related context is the determination of tax liability. (Tax obligations and welfare entitlements are conceptually analogous, particularly when both are determined by income levels.) "Future" tax obligations are usually self-assessed, with litigation arising only when the government disputes a taxpayer's claimed obligation. As with welfare, the expectation of tax obligations has important effects on ex ante behavior; unlike with most other areas of law, the law's effect on behavior tends to be undesirable rather than desirable. For an analysis of the income tax context, see Louis Kaplow, *Accuracy, Complexity, and the Income Tax* (Discussion Paper No. 139, Harvard Law School, Program in Law and Economics 1994).

¹⁸¹ This subsection omits considerations of some other dimensions of legal design. Notably, the ideal level of benefits may depend on accuracy. For example, if accuracy were perfect, one might want disability benefits to provide full compensation (although this conclusion would be affected by taking into account moral hazard). If accuracy were highly imperfect, so that many benefits were paid to individuals not in need, lower benefits would be appropriate. (Since only a fraction of each added dollar of benefits goes to those truly in need, each dollar provides less of a social gain.)

¹⁸² The significance of the issue is indicated by the fact that the Social Security Disability Insurance program alone involves annually over a million claims, payments of more than twenty billion dollars, and public administrative expenses approaching a billion dollars. Social Security Administration, *Annual Statistical Supplement, Social Security Bulletin*, tables 4.A2, 4.A6, 6.C7 (1991).

¹⁸³ Many details will be omitted for simplicity. The same logic would be applicable to

of adjudication involves two types of mistakes, paralleling the form of error in the discussion of liability in Part III. Some disabled individuals will be mistakenly denied benefits, and some individuals who are not disabled will be mistakenly granted benefits.¹⁸⁴ Expenditures to increase accuracy are assumed to reduce both types of errors, so that with more accurate procedures a higher portion of the truly disabled receive benefits, and fewer individuals who are not disabled receive benefits. (This assumption is unconventional, as it is usually assumed that more of both groups would receive benefits; Subsection 3 explains, however, that the conventional view confuses accuracy and the *de facto* burden of proof.)

In order to assess the value of accuracy, consider first the effect on the disabled. Providing more of them benefits has an element of social gain and a cost. The gain corresponds to the social value of a disabled person receiving benefits, rather than being left to other means of support. The cost is that of the funds, which may be supplied by increasing taxes, diverting revenue from other programs, or reducing benefit levels.¹⁸⁵ The net effect should be positive, for otherwise it would be best to eliminate the program.

Second, consider those not disabled. Denying them benefits involves a cost (for, after all, providing payments to those not disabled is not a complete waste of resources)¹⁸⁶ and a gain. The gain here corresponds to

initial determinations or eligibility for continuation of benefits (or to what procedures are required for determinations that take immediate effect but are subject to subsequent reconsideration). Similar reasoning also would govern disputes concerning the level of benefits. Finally, the discussion ignores the important fact that disability is usually a matter of degree, suggesting that the cost of an erroneous decision in some cases (clear ones) may differ greatly from that in others (close ones). See Mashaw, *supra* note 87, at 82–85; Jerry L. Mashaw, Administrative Due Process as Social-Cost Accounting, 9 Hofstra L. Rev. 1423, 1441, 1444–45 (1981).

¹⁸⁴ In both instances, individuals only come before a decision-maker if they apply, whereas in the discussion of accuracy in determining liability in Part III, a victim or the government initiates proceedings. The level of accuracy will affect incentives to apply. As accuracy increases, there is less incentive for one not truly disabled to apply (assuming that they know at the time of application whether they are indeed disabled, which is true in cases of outright fraud but may not be for borderline cases). Greater accuracy increases the likelihood of success for the truly disabled, making application more attractive, but there may be a countervailing effect since more accurate proceedings may cost more in time and resources for applicants. (Application charges or subsidies could be offered to alter these incentives.) See also note 7 *supra* (effect of accuracy on frivolous suits).

¹⁸⁵ One way to reduce benefits is to make it more difficult to qualify for benefits, which Justice Black suggested may be the effect of the Court's ruling in *Goldberg v. Kelly*, 397 U.S. 254, 279 (1970) (dissenting opinion).

¹⁸⁶ An economic approach ordinarily counts gains to all individuals, even when the gains may not be "deserved." Moreover, note that many unqualified individuals may be disabled to some extent or otherwise in need, so dismissing all benefit to their receiving funds would be inappropriate. In any event, the remainder of the argument does not depend on how one treats this issue.

the cost with the truly disabled: reducing expenditures allows one to reduce taxes, increase funding of other programs, or raise benefit levels. The net effect here should be positive as well, for otherwise it would be best to expand the program to include those not disabled.

Finally, there is the cost of providing more accurate procedures themselves. Because this cost must be financed, it also has the character described previously.¹⁸⁷ To determine whether a posited increase in accuracy is desirable, one would determine how much it reduced each type of error and sum the gains and losses just described.

An alternative way to view this cost-benefit calculus is to group all the financial costs and benefits together. More disabled receiving benefits increases the revenue requirement; fewer individuals who are not disabled receiving benefits reduces required revenues; and enhanced accuracy affects administrative costs. Taken together, the net effect on direct costs could be positive or negative. To determine whether increasing accuracy is desirable, one would add to this the benefit of more disabled individuals receiving payments and the cost that fewer individuals not disabled receive payments.

2. Claimants' Incentives to Present Information¹⁸⁸

Subsection 1 indicates that the appropriate level of accuracy is determined by a number of factors. But the claimant will only be concerned with the value of benefits to himself and the costs he incurs in attempting to demonstrate eligibility. Thus, claimants' incentives to provide information to enhance their chance of receiving benefits will be excessive because claimants do not take into account the cost of providing benefits.¹⁸⁹

Mashaw suggests that no social value would be assigned to disability payments to individuals with no disability. See Mashaw, *supra* note 87, at 82–83. Yet, his construct suggests that it is desirable to make awards whenever benefits of the payment are positive, which implicitly ignores the costs of financing benefits (that is, the financial cost of funding the benefits). See *id.* at 83. Taken together, however, these two omissions are largely offsetting—they would be precisely offsetting if the cost of financing benefits just equaled the social value of benefits to those not at all disabled (although generally the cost would be higher since taxation involves administrative and incentive costs).

¹⁸⁷ As suggested in note 184, accuracy may affect incentives to apply for benefits, which will affect administrative costs. If total applications fall, and by a sufficient amount, administrative costs would fall. The analysis in the text does not depend on whether the administrative cost is positive or negative. In addition, the text simplifies in implicitly assuming that all administrative costs are budgetary costs, but they may also involve private expenditures of time and money.

¹⁸⁸ For a formal analysis, making the analogy to private insurance, see Louis Kaplow, *Optimal Insurance Contracts When Establishing the Amount of Losses is Costly* (Discussion Paper No. 122, Harvard Law School, Program in Law and Economics 1993).

¹⁸⁹ Claimants also do not take into account costs of running the system that determines

To illustrate, suppose that a claimant seeks to receive \$10,000 in public benefits. To receive these benefits, the claimant would be willing to spend up to \$9,999. But the social benefit of the claimant being found eligible is always less—and often much less—than this amount. The reason is that paying the \$10,000 in benefits involves an expenditure of public funds, which may require increasing taxes, reducing others' benefits, or cutting other programs. Hence, the social value of the transfer is less than the private value, regardless of whether the claimant is truly eligible or ineligible.¹⁹⁰

In practice, of course, claimants usually would choose to spend far less than \$9,999. For example, if a claimant can spend \$500 to achieve a 95 percent chance of receiving benefits, he surely would not spend an additional \$9,499 to increase the probability to 100 percent, even if this were possible. The argument in the previous paragraph, however, applies equally to incentives at the margin. When a claimant contemplates an additional expenditure to increase the likelihood of receiving benefits, he considers only the cost of the expenditure and the gain from receiving the benefits, and not the social cost of funding the program. Therefore, whatever the amount individuals choose to spend, the amount will be socially excessive unless they are restrained or discouraged from undertaking as much effort in establishing eligibility as they would like. A system designed to determine benefit eligibility should generally constrain to some extent the ability of individuals to expend resources establishing their claims.¹⁹¹

3. Courts and Commentators' Views

Much of the attention to accuracy in the context of benefit eligibility has been with regard to the Supreme Court's approach to procedural due

eligibility, aside from the costs they incur themselves. To simplify the exposition, it is useful to think of these costs as part of the cost of providing the benefits.

¹⁹⁰ If the claimant is truly eligible, the social value of providing the benefits may be high, but it is still the case that there is a nontrivial social cost in providing the funds. If the claimant is ineligible, the transfer—taking into account the benefit to the claimant and the cost of providing the funds—is of negative social value (assuming that the program makes ineligible those individuals to whom it would be undesirable to provide a transfer), and the resource cost incurred in claiming the benefits is an additional social cost.

¹⁹¹ How much they should be constrained is an empirical matter. There are other complications that could favor a different outlook. For example, claimants may have resource constraints that make it impossible for them to finance even the socially appropriate investment in establishing their claim, in which case subsidies may be desirable. The point would remain that the amount of assistance that should be provided is less than what a claimant would be willing to pay if not under a resource constraint (and, a fortiori, less than the amount they would wish to have paid on their behalf at no cost to themselves).

process that was first articulated precisely in *Mathews v. Eldridge*.¹⁹² “[I]dentification of the specific dictates of due process generally requires consideration of three distinct factors: First, the private interest that will be affected by the official action; second, the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or substitute procedural safeguards; and finally, the Government’s interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.”¹⁹³ Much subsequent criticism has focused either on whether the Court in this case or others applied this analysis correctly¹⁹⁴ or on the appropriateness of a cost-benefit framework.¹⁹⁵ The present discussion will instead concern whether the cost-benefit framework is properly conceived.

As an initial matter, the Court’s formulation is confused. The third element—the government’s interest—includes everything and tells nothing. Presumably, even the first two elements are part of the government’s interest.¹⁹⁶ And, if one views the government’s interest narrowly as involving only its selfish interest, the statement is unilluminating because it says nothing about how this interest is affected by changing procedures. The second element is somewhat puzzling, because it does not explain why the risk of an erroneous deprivation is relevant *in addition to* the value of a procedural change. Implicitly, it is assumed that reducing error from 10 percent to 8 percent is more important than reducing it from 4 percent to 2 percent, but no reason is given.¹⁹⁷ Even the first element is

¹⁹² 424 U.S. 319 (1976). For subsequent illustrative applications, see *Connecticut v. Doe*, 111 S. Ct. 2105 (1991); *Lassiter v. Dept. of Social Services*, 452 U.S. 18 (1981).

¹⁹³ 424 U.S. at 335.

¹⁹⁴ Most critics suggest that in applying this test courts are too stingy to claimants because they fail to appreciate the significance of the private deprivation or the value of procedural safeguards and they overstate the government’s interest. See, for example, Jerry L. Mashaw, *The Supreme Court’s Due Process Calculus for Administrative Adjudication in Mathews v. Eldridge: Three Factors in Search of a Theory of Value*, 44 U. Chi. L. Rev. 28, 37–46 (1976). Some, however, have suggested that courts are all too willing to impose additional procedural requirements. See Henry J. Friendly, “Some Kind of Hearing,” 123 U. Pa. L. Rev. 1267 (1975).

¹⁹⁵ In particular, the suggestion is that a crass utilitarian test overlooks other values, notably dignitary values. See, for example, Owen M. Fiss, *Reason in All Its Splendor*, 56 Brooklyn L. Rev. 789 (1990); Mashaw, *supra* note 87, at 88–97; Mashaw, *supra* note 194, at 46–59; Jerry L. Mashaw, *Administrative Due Process: The Quest for a Dignitary Theory*, 61 B.U.L. Rev. 885 (1981). Some of these other values will be addressed in Section VC.

¹⁹⁶ See, for example, Mashaw, *supra* note 87, at 81.

¹⁹⁷ One might imagine that the errors eliminated when moving from 10 percent to 8 percent would, on average, be more severe errors than when moving from 4 percent to 2 percent since when the error rate is low only more modest errors may remain. See note 183 *supra*.

problematic, as it refers to the private interest affected without specifying whether it is meant to include the interest of ineligible applicants or only eligible applicants.

A generous interpretation of the Court's discussion¹⁹⁸ suggests that the first factor refers to the interest of the disabled in receiving benefits, the second to the reduction of false negatives, and the third to costs of the added procedure plus something else, presumably involving the cost of paying benefits to the ineligible. (To be complete, this latter component would have to include the effect on false positives, the value of benefits to the ineligible, and the costs of providing benefits.) Even so, this interpretation is not really a formulation. Rather, it is a heterogeneous listing of relevant factors. There is little care either in specifying each relevant component or in ruling out irrelevant ones. And it does not attempt to indicate how the components fit together.¹⁹⁹

These limitations of the Court's test need not create significant problems as long as those who apply the test take it to represent a more coherent formulation. It does not appear, however, that clear thinking always results. In particular, there seems to be systematic confusion about the relationship between accuracy and the burden of proof, a topic explored in Subsection III B1.

This confusion is reflected in the Court's third factor, which is generally seen as suggesting that government functions will be adversely affected, even aside from the added costs incurred in funding additional procedural safeguards. This seems to assume that the added procedure will involve an increase in the rate at which benefits are paid to individuals who are not disabled.²⁰⁰ Procedures that increase accuracy in isolation do not produce this sort of effect: more accuracy would reduce the rate of mistaken grants of benefits. Instead, this pattern of results—a higher rate of granting benefits to both the disabled and those who are not disabled—is what follows from a reduction in the implicit burden of proof. In other words, the Court and subsequent interpreters and commentators implic-

¹⁹⁸ See 424 U.S. at 339–43.

¹⁹⁹ The Court simply announces the test, with little explanation beyond a citation to *Goldberg v. Kelly*, 397 U.S. 254, 263–71 (1970), which articulates no test but does discuss various factors leading to the Court's conclusion in that case. The interpretation must be made from the Court's application of each component to the facts of *Mathews v. Eldridge*, *Goldberg v. Kelly*, and other cases. In many instances, this leaves little ambiguity over what is included, but the reader often must speculate as to the reasons for inclusion, the relationship among the factors (including their weight), and whether elements not mentioned are excluded or simply not present in the case before the Court.

²⁰⁰ The only alternative explanation would be that the Court had in mind the cost of paying for benefits of deserving individuals, but it seems clear from the context and subsequent applications that this is not the case, at least not entirely.

itly have in mind procedures that involve a de facto shift in the burden of proof, without regard to their effect on accuracy.²⁰¹

This makes the Court's test and discussion of it easier to understand. The origin is apparent: those who challenge existing procedures are individuals who lose under them, and they will seek an additional procedure that will improve their circumstances.²⁰² But there remains the problem of justification. Is there a constitutional requirement that burdens of proof not be too high in determining benefit eligibility? If not, what is the basis for demanding a reduction in the burden of proof? The remaining discussion will not answer these questions but, rather, suggest that, whatever the basis for inquiring into the burden of proof, the conventional approach seems to differ greatly from one designed to address this issue.

First, a necessary element of a challenge to an existing proof burden as inadequate would be an inquiry into the height of the current burden of proof. Perhaps the previously criticized second component—referring in part to the risk of erroneous deprivation—embodies this inquiry. But then it is unclear why the inquiry would not be at an end. If the burden is too high, the complaining party wins; if not, he loses.

Instead, one may understand this as a sort of cost-benefit analysis about the burden of proof.²⁰³ The burden of proof should be lowered only

²⁰¹ For example, Mashaw argues that "[a]ny attempt to eliminate either type of error will induce errors of the opposite type. Hence, a move in the direction of eliminating false negatives—for example, by relaxing standards of proof with respect to some criterion of eligibility—will induce some additional number of false positives." Mashaw, *supra* note 87, at 84–85 (emphasis added). This suggests, remarkably, that improving accuracy *per se*—as by obtaining additional medical tests, increasing training of decision-makers, investigating further the applicant's ability to perform particular tasks—is impossible. (It is apparent from other parts of Mashaw's book that he does not believe this.)

Further evidence that the burden of proof is the core concern is that those who oppose more procedures often do so because of the increase in mistaken grants of benefits, whereas those who support the procedures may dispute the magnitude of the effect but usually do not dispute the direction. Nor do they propose, say, adding procedural protections for claimants while simultaneously increasing the explicit burden of proof so as to keep the likelihood of mistaken grants of benefits constant. (Obviously, such a rise in the burden of proof would result in more mistaken benefit denials than under the proposed added procedures, although less than under the status quo *if the proposed procedures were indeed more accurate*.) Perhaps the assumption that procedures inevitably reduce the claimant's de facto burden of proof has seemed natural because the Supreme Court cases generating much of the literature have involved terminations of benefits, where the claimant was demanding a pretermination hearing. If such a hearing were required, it is generally presumed that there would be added delay before all terminations, which would be pro-claimant in the interim (assuming that mistakenly paid benefits often could not be recovered). See, for example, Mashaw, *supra* note 183, at 1435–36.

²⁰² Even this need not be true. A different procedure may be adverse to claimants in general, but for a claimant who has already lost under existing rules, it is desirable to have a second chance, even if the new game is a more challenging one.

²⁰³ In fact, the Mathews v. Eldridge balancing approach is used in disputes about standards of proof in civil cases. See Laurence H. Tribe, *American Constitutional Law* 740–43 (2d ed. 1988).

if two conditions simultaneously hold: the burden is “high,” and lowering it is justified on cost-benefit terms. Alternatively, the idea might be that some combination of excessive height in the proof burden and sufficiently good marks on a cost-benefit test for its reduction warrant finding for the claimant. While one can understand such an approach, it seems to diverge substantially from much of the discussion and application of the Court’s test.²⁰⁴ Yet another formulation would ignore the aspect of the second component referring to the rate of error and simply understand the formulation as stating that there should be a cost-benefit analysis of the burden of proof in such settings. If the burden is inefficiently high, it must be reduced.²⁰⁵

A more puzzling, second question is: if the problem is an excessive burden of proof, why should it be reduced by adopting some particular procedure, whether a hearing, providing a lawyer to the applicant, or whatever? As emphasized in Subsection III B1, there are many ways to alter a proof burden. In the benefit determination context, it may be feasible to do this for free. One could instruct those who determine eligibility (whether case workers, administrative law judges, or those providing internal review of appeals) to require somewhat less confidence in eligibility as a prerequisite to granting benefits. By reviewing the rate at which applications are granted or by other means, such a shift could probably be implemented.²⁰⁶ Even if a direct shift in the burden of proof were infeasible, the question would not be whether a given procedure accomplishes a cost-beneficial reduction in the burden of proof but, rather, what procedure accomplishes a given reduction in the burden of proof at the lowest cost. One could then ask whether implementing *that* procedure would achieve a cost-beneficial reduction in the proof burden.

Thus, if the burden of proof is the object of the inquiry, a rather different approach would be appropriate. Yet there are reasons for caring about the effects of particular procedures, even those that may not involve the cheapest way to reduce the proof burden. An obvious reason is that procedures affect accuracy.²⁰⁷ Thus, providing an opportunity for claim-

²⁰⁴ Most notably, one does not find extensive analyses of what “too high” a burden of proof may mean.

²⁰⁵ If it is too low, the government is always free to amend statutes, regulations, or procedures.

²⁰⁶ Presumably, internal administrative guidelines and procedures already pay much attention to the implicit, *de facto* burden of proof. Those granting or denying eligibility may get various signals and sanctions that induce them to behave differently. There is no obvious reason that a denial rate that was a few percentage points higher or lower would involve much difference in administrative costs.

²⁰⁷ Other reasons are discussed in Section VC.

ants to explain their situation and respond to arguments may improve decision-makers' knowledge of which claimants are eligible. But the proper analysis of whether a gain in accuracy is desirable (explained in Subsection 1) is rather different from that for determining how high the proof burden should be.

The analysis in Subsection IIIB1 emphasized that it is conceptually useful and typically feasible as a practical matter to separate evaluations of accuracy and of changes in the implicit burden of proof. It appears that courts and commentators that have addressed procedural requirements for determining benefit eligibility care about both and have confused the two inquiries.²⁰⁸ The result is that the objectives of the analysis are not clear, and the analysis of both accuracy and burdens of proof suffers. Thus, results concerning particular procedures viewed in isolation may be mistaken and, even if they are narrowly correct, they may not involve the best procedural design compared with alternative ways to affect the burden of proof and accuracy.²⁰⁹

When it is proposed, say, that there be a more formal hearing, it should separately be asked whether the hearing would produce a cost-justified increase in accuracy and whether it would result in a desirable reduction in the burden of proof. If both answers are affirmative or negative, the outcome is clear. If only the former answer is affirmative, it would be desirable to adopt the procedure but also to raise the proof burden by other means.²¹⁰ If only the latter answer is affirmative, it is probably best not to adopt the procedure but, rather, to reduce the proof burden in some other manner.

²⁰⁸ The best evidence for this is that, in concrete applications, a procedure's effects are almost uniformly described in a manner that implies a shift in the burden of proof, with no necessary effect on accuracy in either direction, while general discussions of the desirability of additional procedures often emphasize accuracy and make no mention of how procedures affect the burden of proof.

²⁰⁹ In part the problem reflects the limits of designing a legal system piecemeal, through isolated disputes about the propriety of particular procedures. Also, government officials may anticipate that courts will be inflexible. For example, if a court requires new procedure *X* today, an agency might imagine eliminating *X* tomorrow and substituting *Y*, which it believes accomplishes the objective in a more efficient manner. The question is whether the court's order is interpreted literally to require *X*, or only to require certain sorts of results, so that it may require *X* in the absence of *Y* but would not require *X* if *Y* were present. Since courts' statements of objectives are so obscure, this is difficult to determine.

²¹⁰ Such an outcome may occur presently. When an agency is forced to adopt an added procedure, particularly if it has a limited budget, it may be induced to adjust its internal operations in a manner that leads decision-makers to be more stingy with regard to a given set of information. See Mashaw, *supra* note 183, at 1447–48; notes 137 and 206 *supra*. See also note 185 *supra* (agencies may be more reluctant to grant initial eligibility if termination is made more difficult).

B. Adjudication about Past Acts That Governs Future Conduct

1. The Effect of Remedies Specifying Future Conduct on Prior Conduct

Often, adjudication combines elements of past behavior and future conduct. For example, prior conduct violating another's rights may be enjoined or one who commits a crime may be imprisoned. In such instances, the analysis of Parts II and III continues to be applicable, as the anticipation of an injunction or imprisonment affects individuals' *ex ante* decisions about how to behave. But such *ex ante* analysis may be only part of the story. Injunctions and imprisonment dictate future behavior as well. In contrast, when sanctions are monetary (damage awards and fines), individuals are penalized for their prior behavior but left to their own decisions with regard to the future.²¹¹

This difference, however, may be overstated in the case of injunctions, for often the enjoined party has the option of contracting with victims to remove the restraint imposed by the injunction.²¹² Similarly, in the absence of an injunction, a victim might contract with an injurer to have the latter refrain from her harmful conduct in the future. Thus, it is familiar that the difference between property rules (entitlements enforced by injunction) and liability rules (entitlements enforced by damages) is less than may appear. In the limit, when bargaining proceeds without any cost or impediment, they are the same.²¹³ In such instances, any injunction can be seen as having a monetary equivalent, so the analysis of Part II concerning accuracy in assessing damages would be applicable.

Nonetheless, it is often the case that injunctions remain in place because bargaining is imperfect or is prohibited. The problem of numerous dispersed victims is the most familiar example of the former. For the latter, one might consider civil rights injunctions or a declaration that a

²¹¹ The prior adjudication may still affect future behavior. First, individuals' information may change. (One may not have known previously that an act was illegal, but for the future one knows and is likely to remember.) Second, if future penalties depend on prior violations, one's incentives in the future will differ. (A proper analysis would take account of the effect on future opportunities in determining the effect of a first sanction on an individual's *ex ante* behavior. See, for example, A. Mitchell Polinsky & Daniel L. Rubinfeld, *A Model of Optimal Fines for Repeat Offenders*, 46 J. Pub. Econ. 291 (1991).)

²¹² Also, an enjoined party may have the option of violating the injunction and paying a penalty.

²¹³ See, for example, Guido Calabresi & A. Douglas Melamed, *Property Rights, Liability Rules, and Inalienability: One View of the Cathedral*, 85 Harv. L. Rev. 1089 (1972); R. H. Coase, *The Problem of Social Cost*, 3 J. Law & Econ. 1 (1960).

statute is unconstitutional.²¹⁴ Also, one cannot be released early from prison upon paying a fee.²¹⁵

2. The Value of Accuracy with Regard to Future Conduct

An analysis of the benefit of accuracy with regard to future rights and obligations is more similar to conventional cost-benefit analysis, as of a public works project or of eligibility for future entitlements, than to that in Parts II and III. Compare, for example, crafting the right injunction for operating a factory that pollutes with determining the best location for a road. In each case, one can always undertake more study in the hope of improving the decision, but perfect accuracy is neither feasible nor desirable, due to its cost.²¹⁶

When adjudication both creates incentives for *ex ante* behavior and affects future conduct, accuracy tends to be more valuable. Consider imprisonment, where the benefit with regard to future conduct is incapacitation, which prevents dangerous individuals from causing further harm to the community.²¹⁷ Note that the benefit of incapacitation does not depend directly on whether individuals have accurate *ex ante* knowledge.²¹⁸ The benefit of accuracy concerns whether those incapacitated are indeed those who are dangerous. If one increases accuracy, there are two effects. More guilty, presumed dangerous, individuals are incapacitated, which is an advantage from this perspective. In addition, fewer innocent individuals are incapacitated. Incapacitating the innocent is of no independent benefit (and is even likely to be harmful, in addition to

²¹⁴ There is some room for bargaining even in this latter area. For a civil rights injunction, the injunction may be avoidable if, prior to a court order, the case is settled with a monetary payment. (This may have to be court approved, however.) For problems of constitutionality, if the benefits of a different approach are sufficiently great, the constitution may be amended. Needless to say, in both contexts there often are sufficient impediments to bargaining that a court's injunction will remain in effect even if it is not in the aggregate interest of affected parties.

²¹⁵ In some instances, there may be a *de facto* choice between monetary payments and imprisonment, as when plea bargaining involves an agreement from a prosecutor to recommend a particular type of sentence.

²¹⁶ One way injunctions sometimes account for error is to provide for modification in light of future information.

²¹⁷ A more extreme situation arises with capital punishment, for which the Supreme Court has emphasized "the Eighth Amendment's heightened 'need for reliability in the determination that death is the appropriate punishment.'" *Caldwell v. Mississippi*, 472 U.S. 320, 340 (1985) (quoting *Woodson v. North Carolina*, 428 U.S. 280, 305 (1976) (plurality opinion)).

²¹⁸ *Ex ante* knowledge may be relevant. For example, if an individual would not have violated the law if she knew that the act was illegal, incapacitation for its own sake would be unnecessary.

the deprivation of liberty, if prison teaches the uninitiated how to be criminals). It was previously noted that the high social cost of imprisonment as a sanction justifies greater accuracy than otherwise.²¹⁹ Considering incapacitation in addition to deterrence and sanction costs suggests that accuracy in determining imprisonment should be higher still.²²⁰

3. Parties' Incentives to Present Information in Adjudication

In this context, as in the others examined in this article, there is no particular reason to believe that parties' incentives to present information are socially optimal.²²¹ Parties gain from adjustments that are purely redistributive or even those that decrease value. (The discussion in Subsection A2 of claimants' excessive incentives to demonstrate eligibility for public benefits illustrates this problem.)²²² Parties' private gains from presenting information may thus systematically exceed social benefits. An exception²²³ may arise when there are third parties that are unrepresented or ineffectively represented.²²⁴ The problem, however, is not that those parties effectively represented will present too little information; their incentives remain excessive. Rather, information that it is not in their interest to present may remain unavailable to the decision-maker.

²¹⁹ See Subsection IIIA2b *supra*.

²²⁰ The argument is applicable to determining not only who is imprisoned but also the duration of the term. The reason is that, from an incapacitation perspective, it is desirable to imprison the more dangerous for a longer period of time. Thus, for example, even if one is certain as to whether an individual committed murder, knowing whether it was premeditated or committed in the heat of passion may be relevant to assessing dangerousness for purposes of incapacitation. (Individuals who kill in the heat of passion have lower recidivism rates.)

²²¹ It is useful to contrast the manner in which information is collected to design public projects and to determine the content of injunctions. The adversary system is obviously much more central in the latter context, although the use of public hearings or other opportunities to comment on proposed action introduces elements of the adversary system into the former.

²²² Expenditures by truly eligible claimants may be socially desirable even though their objective is redistributive because this particular redistribution is socially favored, but the analysis suggested that, as in other distributive settings, individuals ignore the cost of making the payment they will receive. (Expenditures by those who are ineligible involve, as described in note 190, both undesirable redistribution and a waste of resources.)

²²³ In addition, the exception discussed in Subsection IIID2b (that individuals do not take into account the full social cost of costly sanctions) remains applicable.

²²⁴ Class actions and intervention are devices designed to address these problems, but familiar agency problems render them imperfect. See, for example, John C. Coffee, Jr., *The Unfaithful Champion: The Plaintiff as Monitor in Shareholder Litigation*, 48 L. & Contemp. Probs. 5 (Summer 1985); Jonathan R. Macey & Geoffrey P. Miller, *The Plaintiffs' Attorney's Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform*, 58 U. Chi. L. Rev. 1 (1991).

V. OTHER VALUES OF ACCURACY

This article presents an economic analysis of the value of accuracy in adjudication. The primary focus throughout is on how accuracy affects the central instrumental purposes of the law. Therefore, Parts II and III, which discuss accuracy in determining damages and liability, emphasize how accuracy is relevant to the effectiveness of the law in controlling behavior. Part IV, on accuracy in establishing future rights and obligations, concerns the extent to which the law's purposes in directing future events are fulfilled. The analysis also emphasizes costs of achieving the law's objectives—in particular, administrative costs and social costs arising from the imposition of sanctions.

Yet, much concern expressed about the accuracy of the legal system appears to be addressed to different objectives. Thus, accuracy may be valued because it advances fairness and justice, one cares about mistakenly sanctioning the innocent *per se*, it is important to the system's legitimacy that adjudication appears to be as accurate as possible, or abuse of power is made more difficult. Also, some procedures that affect the accuracy of adjudication may be valued in themselves, perhaps because they enhance individuals' participation in processes that will affect their lives.

While there is much that has been and could be said about these issues, little attempt will be made to address them here. First, in most instances, the nature of the value at stake is not very well defined, and its grounding is uncertain.²²⁵ This makes application in concrete contexts and evaluation of trade-offs quite difficult. Second, the type of analysis that would be appropriate may be rather different from that offered here, so it seems useful to undertake further explorations separately.

Nonetheless, the analysis in Parts II–IV offers some illumination of many of these other values. This part notes some respects in which these seemingly different values may be implicitly included in the economic analysis. It also indicates how the economic analysis may illuminate the extent to which other values are implicated in the contexts that have been examined here. These remarks will not be comprehensive. Nor can they be definitive with respect to the topics addressed because many of these values can be articulated in so many ways that an argument relevant to one version may be inapplicable to another.

²²⁵ For example, Mashaw refers to his own statement of the dignitary values of due process as follows: "The values that fit our intuitions are vague at the margins and potentially contradictory at the core." See Mashaw, *supra* note 195, at 899.

A. *Accuracy in the Assessment of Damages*

1. Fairness and the Accuracy of Compensation

With regard to accuracy in assessing damages, one might be concerned with the fairness of the outcome in addition to how the accuracy of adjudication affects behavior. Thus, if a defendant is required to pay more than the harm she caused, this may be seen as an unjust burden. Or if she pays less than harm caused, this may be seen as less than she deserves to pay. Similarly, if a plaintiff collects too little, corrective justice may not be served. If he collects too much, it may be viewed as an undeserved windfall.

Examine this claim first in the simplest setting examined in Part II, in which individuals are risk-neutral and potential injurers at the time they act do not know their true level of harm (and cannot reasonably determine it in advance). Then, arguably, the analysis suggests that the result is fair. Both an accurate system (damages equal to actual harm) and an inaccurate one (damages equal to average harm) result in injurers being obligated for the same *expected* payment. Moreover, if they are risk-neutral, they value these two outcomes identically. Thus, an *ex ante* notion of fairness may be appropriate, and it would not require greater accuracy. (An analogous argument could be made for plaintiffs.)²²⁶

When individuals are risk-averse, it is no longer true that they are indifferent between damages equal to average harm and actual harm. For plaintiffs, Subsection II.A.3 noted that the possibilities of under- and overcompensation are costly because of the risk that is imposed. As a result, greater accuracy would be justified on this account. The question then becomes whether fairness demands even more accuracy than this. That is, does some fairness norm require that the system attach more weight to mistakes in compensating plaintiffs than would be attached by

²²⁶ Accuracy may be relevant for *ex ante* fairness to plaintiffs if some suffer systematically more harm from accidents than others, so that, *ex ante*, they can expect to be undercompensated. Individuals whose damages, measured in dollars (as the legal system does even when the injuries are nonpecuniary), will usually be above average are typically the wealthy, who have more skills (and thus lose more wages when disabled), attach higher economic value to nonpecuniary losses, and have more expensive property to be damaged (a Cadillac rather than a compact car). First-party insurance would eliminate variance in awards, but the wealthy would pay higher premiums to the extent damages do not equal actual harm. (For example, auto insurance premiums for a Cadillac are higher than for a compact car.) Of course, adjustments elsewhere, as in the tax system, could offset such an effect. See note 238 *infra*. Whether the disparate effect is correctable and whether it is seen as unjust will likely depend on its source and who is affected.

the plaintiffs themselves? If not, the economic analysis subsumes this fairness concern.

For defendants, risk aversion indicates that less accuracy than otherwise is warranted because accuracy imposes risk. Does fairness demand that defendants bear such risk, when it is costly to them and they could not reasonably foresee the difference? If not, fairness would not call for a prescription different from that in the economic analysis.²²⁷

Another aspect of fairness concerns notice. If, indeed, defendants could reasonably foresee average harm but not actual harm, do they deserve to pay for actual harm, even when it is unusually high or low? If not, an accurate system may be more objectionable than an inaccurate one.²²⁸

Now, consider the case in which injurers do know the actual harm they would cause in advance. The economic analysis indicates that accuracy is desirable in this case if it is not too costly. Does fairness demand accuracy at any cost? If not, how much accuracy does fairness demand? In particular, is there a reason to provide more accuracy than the parties would demand if they could agree on the level of accuracy in advance? If not, then fairness would again be subsumed in the economic analysis.²²⁹

2. Ex Post Incentives to Present Information in Adjudication

The analysis in Section IIC indicates that in many settings parties may have an excessive incentive to present information in adjudication. Thus, the economic analysis suggests that the system should be designed in a manner that either reduces these incentives or prevents parties from acting as they wish (as by limiting discovery or the number of experts who may testify). Are such limitations unfair? Would they frustrate the value of allowing parties the ability to participate fully in determining their fate?

A prerequisite to answering such questions is specification of the im-

²²⁷ Observe that, if fairness does require that defendants bear such risks, liability insurance should be prohibited.

²²⁸ Many concerned with distributive justice see eliminating wealth disparities caused by fortuity to be a central concern. Compare Kaplow & Shavell, *Legal Advice*, *supra* note 5, at 603 (fairness of limiting advice about information to present in litigation when individuals do not know at the time they decide how to act what the advice would be).

²²⁹ While the language of the discussion involved cost-benefit analysis rather than hypothetical bargaining, it is familiar that informed parties would contract to an efficient result. The particular application is that dispute resolution is indeed often specified by contract. If fairness demands an outcome inconsistent with what parties would choose, such contractual provisions should not be enforced.

PLICIT notion of full participation. If the system offers both sides one trial day to present evidence, is participation full, excessive, or inadequate? Can one specify full participation independently of the instrumental values that participation serves? One could ask individuals how much they value a day in court and then ask them how much they value two or three days. A result differing from that in the economic analysis would arise only if individuals attached values unrelated to the outcome participation would produce. (For example, if they preferred three days to one, at a cost of, say, \$1,000 per day in lawyers' fees, even if the outcome would be the same, then limiting them to one day might be seen as undesirable.)

Parties usually do prefer to present more evidence because they believe it will advance their case. Similarly, they prefer that their opponent present less. One might ask what joint arrangement they would find best. The answer, in most instances, is zero trial days; most cases settle for precisely this reason.²³⁰ Moreover, one suspects that, when a case fails to settle, this may be because of an inability to reach agreement on amounts of money rather than because one party attaches substantial independent value to the days in court.²³¹

The value of presenting information in court should be viewed *ex ante*. The economic argument that incentives may be excessive concerns the difference between the effect of further information on *ex post* results and on *ex ante* behavior. In the simple case in which *ex ante* behavior is totally unaffected by the prospect of greater accuracy, achieved by allowing additional presentation of information in adjudication, one could ask whether fairness or process values call for a different outcome. Again, the contractual argument can be made: the efficiency analysis suggests what informed parties would have agreed to *ex ante*.²³² This

²³⁰ Relatedly, there are reports of satisfaction with alternative dispute resolution programs that provide an appearance before a neutral decision-maker that falls significantly short of a full trial. See, for example, Steering Committee Report, American Law Institute Study on Paths to a "Better Way": Litigation, Alternatives, and Accommodation, 1989 Duke L. J. 811, 817.

²³¹ In most contexts, parties attach negative value to time in court, both because of direct litigation costs and because of the time and anguish involved. There are, no doubt, exceptions, such as in the libel context where the plaintiff or defendant may wish a public pronouncement. (Even there, a settlement in which the other party agrees to a public apology is often viewed as an adequate substitute, suggesting that it is the outcome—here, a nonpecuniary one—rather than the days in court themselves that are valued.)

²³² That is, if accuracy *ex post* is inefficient, but *ex post* incentives would be such that one or both parties would nonetheless have an incentive to undertake expenditures to present information, the parties would wish to bind themselves in advance not to pursue

leaves two possibilities: either there exist values that would call for nullifying an *ex ante* agreement to limit presentation of information in adjudication²³³ or, if parties attach value on their day in court for its own sake, the economic analysis presented here omits a component.

B. Accuracy in the Determination of Liability

1. The Extent to Which the Analysis Incorporates a Concern for Mistakenly Sanctioning the Innocent

The most obvious independent concern in the context of accuracy in the determination of liability is a concern for mistakenly sanctioning the innocent. Recall, however, that the economic analysis takes into account that sanctions themselves may be costly. Thus, the cost of mistakenly imposing sanctions on the innocent is a central part of the analysis in Part III. To be independently concerned with sanctioning the innocent suggests that in some sense the sanction cost, as measured in an economic analysis, involves an undervaluation.

Consider first the indirect cost of sanctioning the innocent—its effects on their behavior. The analysis suggested that individuals would, on account of the possibility of mistakes, be more likely to commit harmful acts²³⁴ and be more likely to be deterred from committing harmless acts that subjected them to the risk of sanctions.²³⁵ Thus, the analysis fully incorporates the adverse effects of mistakes on the behavior of individuals who are truly innocent. The value placed on this effect is individuals' private valuations of the acts deterred.²³⁶ As long as there are no externalities, there is no obvious reason to depart from such an analysis.²³⁷

Now consider the direct costs of sanctioning the innocent—that is, the costs of the innocent bearing the sanctions themselves, rather than any effect the prospect of sanctions may have on behavior. When sanctions are monetary and individuals are risk-neutral, the analysis treats sanc-

this option. See Subsection IVA2 *supra*; Kaplow, *supra* note 188. Parties in actual contracts who specify simple forms of alternative dispute resolution or who provide for liquidated damages in effect do just this. See Subsection IIB4a *supra*.

²³³ See, for example, Owen M. Fiss, *Against Settlement*, 93 Yale L. J. 1073 (1984).

²³⁴ See Subsection IIIA1 *supra*.

²³⁵ See Subsection IIIA3 *supra*.

²³⁶ In the economic calculus, the cost of deterring a harmless act, for example, is the individual's benefit of the act minus the harm of zero.

²³⁷ If there are externalities, such as benefits to third parties, there is a reason to encourage such acts without regard to the problem of inaccuracy.

tions as socially costless because they involve mere transfers of money. Thus, sanctioning the innocent receives no independent weight. Should this be otherwise? When individuals are risk-neutral, by definition they do not attach any cost to the risky aspect of sanctions. All they care about is the expected value. That individuals committing a harmless act have a positive expected sanction cost does deter desirable behavior, a consequence already noted and taken into account. This leaves the possibility of an argument that it is unfair that individuals who are not deterred must make payments, essentially an argument that there is an unfair tax. Such arguments are much akin to arguments about ex post accuracy of compensation when individuals are risk-neutral, already examined in Subsection A1.²³⁸

When sanctions are monetary and individuals are risk-averse, the riskiness of sanctions is itself a cost. This cost is incorporated in the economic analysis, again with a valuation corresponding to the value innocent individuals themselves would place on it.²³⁹ When sanctions are nonmonetary, as in the case of imprisonment, the social costs of sanctions are viewed as including the loss of liberty.²⁴⁰

In all these cases, therefore, the economic analysis includes the cost of sanctions borne by innocent individuals. Thus, an independent concern for sanctioning the innocent suggests that there is a harm independent of the harm to those who are sanctioned. In addition, as noted, such a concern must be independent of the effect of the anticipation of mistakes in distorting behavior. Finally, if it is a harm unique to sanctioning the innocent, it suggests that it must not be a general cost of sanctions, but one unique to their misapplication.²⁴¹

If one believes that there is such a concern,²⁴² the analysis would be affected in the manner described in the following subsection. The point

²³⁸ There may also be a concern with the effect on the distribution of wealth, if innocent behavior is disproportionately undertaken by, say, low-income individuals (which, for the system as a whole, seems implausible). Compare note 226 *supra*. But it is difficult to maintain arguments that such an effect should justify inefficient redesign of the legal system. See Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, J. Legal Stud. (in press 1994).

²³⁹ That is, the cost of riskiness in sanctions is measured by the amount individuals would be willing to pay, above an actuarially fair amount, for insurance against the risk.

²⁴⁰ Actually, the loss of liberty is used as a shorthand to include any costs to the individual. The argument is unaffected by being more complete, including effects on one's reputation or other long-term effects arising from a mistaken conviction.

²⁴¹ The fear of government abuse of power, see Subsection C3 *infra*, is one possibility.

²⁴² For a discussion suggesting that there is, see Ronald Dworkin, *A Matter of Principle* 72–92 (1985).

of this subsection is merely that many, even if not all, of the undesirable aspects of mistakenly sanctioning the innocent are incorporated in the economic analysis.

2. How a Greater Concern for Sanctioning the Innocent Would Affect the Analysis

If sanctioning the innocent should be given more weight, the analysis would be qualitatively the same. One would consider the same trade-offs, and changes in various components of the system—sanctions, enforcement effort, accuracy, and the burden of proof—would have the same types of effects. The only difference would be that effects with regard to sanctioning the innocent would be weighted more heavily in designing the legal system. Thus, more accuracy than suggested in Subsection IIIA2b and a higher burden of proof than suggested in Subsection IIIB2 may be appropriate.

A common intuition is that concern for sanctioning the innocent also warrants the use of lower sanctions than otherwise. The analysis in Part III,²⁴³ however, suggests otherwise. First, for any given level of deterrence one hopes to achieve,²⁴⁴ greater enforcement effort will be required if sanctions are lower. But raising enforcement effort means that more individuals will be subject to sanctions, so more mistakes will be made. It was argued that, as a first approximation, the effects would be off-setting, so that the expected costs of sanctioning the innocent would be unaffected.

Second, the analysis suggested that relying more on enforcement effort than higher sanctions to achieve deterrence would be counterproductive. When greater enforcement effort is required, more cases are processed (more audits, more prosecutions). This means that each increment in accuracy becomes more expensive. The result is that the appropriate level of accuracy is lower, producing more mistakes.²⁴⁵

²⁴³ See Subsections IIIA1b and IIIA2a *supra*. Details and qualifications noted previously will not be repeated here.

²⁴⁴ The costs of mistakenly sanctioning the innocent may be a reason to reduce the target level of deterrence. (But mistakes also directly reduce deterrence, as explained in Subsection IIIA1, so whether an additional reduction is justified is not obvious.) But, whatever is ultimately deemed to be the appropriate level of deterrence, there remains the possibility of achieving it with high sanctions and low enforcement effort or with low sanctions and high enforcement effort.

²⁴⁵ One may not be *required* to reduce accuracy. But, whatever one spends to achieve accuracy with a high enforcement effort, low sanction strategy, the same expenditure under a low enforcement effort, high sanction strategy would achieve greater accuracy because the same resources would be concentrated on fewer cases.

C. *Accuracy in Establishing Future Rights and Obligations*²⁴⁶

1. The Right to Appear

In the context of determining future entitlements, a commonly expressed objection to employing an economic analysis to determine appropriate procedures is that the process value of allowing individuals to be heard is ignored.²⁴⁷ There remains some question, however, about the

²⁴⁶ The issues examined in this section arise in other contexts as well but have received greater emphasis with respect to government decisions about entitlements, so for convenience they are considered here.

Also, no attempt is made to be exhaustive with regard to possible process values. One not considered here that has received substantial attention is the importance of an independent adjudicator. See, for example, Martin H. Redish & Lawrence C. Marshall, *Adjudicatory Independence and the Values of Procedural Due Process*, 95 Yale L. J. 455 (1986). They offer both instrumental (see *id.* at 476–81) and noninstrumental (see *id.* at 482–91) reasons for this, although the latter do not emerge as having clearly independent weight: “Our analysis will demonstrate that most of the proposed values are inherently tied to the instrumental justification, and cannot be separated from the individual’s interest in his entitlement. Because each of them, with one possible exception, is a variation on the instrumental theme, each of them, in the end, requires the same basic procedures for its achievement.” *Id.* at 482–83.

²⁴⁷ See, for example, Mashaw, *supra* note 194, at 49–52 (linking the right to participate personally in the adjudication with individuals’ dignity). Notes in the remainder of this section often will discuss Mashaw’s views, as he has devoted greater effort to illuminating these issues than others. Yet all the discussion is implicitly qualified by the fact that his ultimate policy conclusions in the area of social security disability claims—a context he has examined in great depth—give less weight to other values than one might have expected from his general discussions of them. See Mashaw, *supra* note 87, at 222–27; Mashaw, *supra* note 195, at 887 (while “a dignitary approach to administrative due process has merit . . . [t]he value . . . is not, however, that it would supplant a positive, instrumentalist, and utilitarian conception of process”; dignitary theory “makes modest ‘absolute’ demands on processes of public decisionmaking”); *id.* at 922, 930 (liberal thought gives only limited determinate content to claims to participation). See also Lance Liebman & Richard B. Stewart, *Bureaucratic Vision*, review of *Bureaucratic Justice*, by Jerry L. Mashaw, 96 Harv. L. Rev. 1952 (1983).

Another prominent exposition of the values of procedural protection aside from their effect on accuracy is Robert S. Summers, *Evaluating and Improving Legal Processes—a Plea for “Process Values,”* 60 Cornell L. Rev. 1 (1974). Summers is explicitly modest in his claims, but the pattern of argument is problematic even for limited conclusions. Summers’ proposed process values are usually defined and justified in a couple of paragraphs. See *id.* at 20–27 (offering ten values). Frequently, his argument involves presenting an instance in which his posited procedure seems attractive. For example, to illustrate the virtue of “procedural fairness,” his example considers an adjudicator who grants advantages to one side not made available to the other. *Id.* at 24–25. Recognizing that this would bias the results, he simply poses the rhetorical question of whether it would be unfair even if it did not. *Id.* at 25. But such a proposition is best tested with an example that does not involve such an effect on results. Thus, consider whether it would be unfair if a plaintiff were given three docket entries for every item he files while the defendant only gets one docket entry. It seems difficult to argue that there is a significant injustice. But if injustice can only arise when there are effects on the outcome, it is hard to see how Summers’ criterion of indepen-

justification for such a process value—in particular, one not already subsumed in the value of accuracy or in setting the burden of proof.²⁴⁸

One suspects that claimants who object to not being heard are those who are, for example, denied benefits.²⁴⁹ If only losers complain, however, one should be suspicious that the complaint is motivated by a concern for the result, and thus an objection to a lack of process may implicitly be an instrumental argument.²⁵⁰ An entirely plausible reason to object to not being heard is that one may believe (perhaps feel certain) that the decision was adverse precisely because the decision-maker was deprived of information one had to offer.²⁵¹ Thus, the decision may have been

dent concern is established. See *id.* at 14 (giving as one of three requirements of a process value that "it is a value that can render a feature of the legal process prizable regardless of its effect, if any, on results"). Even if the example here is unfair to Summers, it does seem proper to ask why he and most other commentators advancing the existence of process values wholly independent of results rely primarily or exclusively on examples with effects on results which are alone sufficient to warrant their preference for the procedure. See *id.* at 46 ("Identification is an acute problem because most process features capable of implementing process values are at the same time capable of serving as means to certain outcomes."). Summers also argues that his position is supported by the fact that "some process values are so taken for granted in our own society." *Id.* at 33. But some procedures may be valued precisely because of their effect on results. Moreover, it is difficult to justify the posited values on grounds of public support and then criticize society for failing to act on them (which suggests that the support is minimal or absent). See *id.* at 39, 42.

²⁴⁸ Compare William Van Alstyne, Cracks in "The New Property": Adjudicative Due Process in the Administrative State, 62 Cornell L. Rev. 445, 475–76 (1977):

To the extent that one finds procedural grossness offensive in food stamp administration, public housing, public employment, etc., it is almost certainly because one believes that access to these things is itself of importance. Mistaken denial, or mistaken termination of such things to those who are in fact eligible is per se the wrong that makes one angry. It belittles the matter to rest the case for minimal procedural fairness only on an extended anxiety in behalf of first amendment interests, which we know in advance will be involved in but a tiny fraction of the cases in which mistakes are made.

²⁴⁹ One does not often hear stories of individuals who win complaining that they did not get their day in court. If there is an independent process value, they would have a claim for reconsideration under proper procedures despite their victory, and if they valued the process significantly, they would choose such a reconsideration, even at some risk of losing.

²⁵⁰ "We all feel that process matters to us irrespective of result. This intuition may, of course, be a delusion. We may be so accustomed to rationalizing demands for improvement in our personal prospects on the purportedly neutral terms of process fairness that we can no longer distinguish between outcome-oriented motives and process-oriented arguments. Thibaut and Walker's experimental work . . . may, after all, merely demonstrate that we generally regard control or the opportunity for personal strategic behavior as the best protection for our substantive concerns." Mashaw, *supra* note 195, at 887 (footnote omitted).

²⁵¹ Information should be construed broadly. For example, if the decision-maker would have observed the applicant's demeanor, and this would be relevant to the decision, then not allowing applicants to appear in person involves a reduction in information.

Mashaw suggests that claimants would perceive a system as unfair if their claims received

inaccurate. Alternatively, one may suspect that the decision-maker would be more favorable when the claimant appears personally, independent of any additional information made available, suggesting a favorable shift in the implicit burden of proof.

To test this, one must consider a hypothetical situation—one probably too far removed from the typical disappointed applicant's mind for him to take seriously—in which the applicant is heard but it is certain that the decision would be unaffected by the hearing. Would individuals value appearing if they knew in advance that they would be ignored or that they would be "heard" but that hearing them could have no effect whatever on the decision?²⁵²

From one perspective, this is simply an empirical question that could be tested directly.²⁵³ There is indirect evidence relevant to how much

less efforts than those of others. See Mashaw, *supra* note 87, at 90; see also *id.* at 90–91 (claimants would be alienated by Kafkaesque procedures). If they won, it would be surprising if they felt this way. If they lost, they would likely perceive it unfair if they suspected that, as a result of inadequate effort, a decision-maker did not obtain information that may have led to a different decision. (Note that, since claimants voluntarily supply self-serving information, one might think they would feel unfairly treated when more rather than less effort was devoted to investigating their claims.)

²⁵² See, for example, Redish & Marshall, *supra* note 246, at 487–88:

This participation only makes sense, though, if the individual harbors some hope of bringing about substantive change in the state agent's action or attitude. Of course, the change in attitude might affect only future cases, not this participant's situation, but nonetheless the focus remains results-oriented. . . . Even as Michelman defines it, participation means "full and frank interchange," and thus focuses on the litigant's opportunity to inform the agent in the hopes of changing her decision. . . . The connection that many have drawn between "control" and "participation" demonstrates the inseparable connection between participation and result efficacy. Imagine a situation in which the state agent announces that his mind is absolutely made up and that he will not reconsider his decision. Does participation at that stage afford any opportunity for "control" over one's own destiny?

For example, Mashaw argues that "participation increases self-respect to the degree that participation gives the participant *control* over the process of decisionmaking." Mashaw, *supra* note 87, at 95. Of course, a claimant would presumably exercise his control to increase the odds of success. Thus, self-respect is said to require allowing claimants to increase their chances of success (apparently irrespective of the merits)—a rather odd argument. Is it that self-respect demands a greater likelihood of success than the merits indicate? Or that, whatever level of success is given a claimant—however high—he should be entitled to increase it personally? (Even if affording this opportunity leads to a change in the initial burden of proof leaving the claimant's chances of success unchanged or reduced?) Mashaw finds this and other participation values persuasive but does not attempt to indicate what they mean or why they should be given any weight. See *id.* at 95–96. Compare Tribe, *supra* note 203, at 666–67 (noting "the special concern about being personally *talked to* about the decision rather than simply being *dealt with*").

²⁵³ For example, one could have two systems, known to produce identical outcomes, but in only one is the applicant heard. By charging differential fees, one could measure the value individuals associate with the procedure. Of course, the experiment may fail to provide much insight because it would be hard to convince the applicants that their being heard

people value such appearances for their own sake. One type of evidence noted previously is the high rate of settlement in most civil litigation.²⁵⁴

would indeed have no effect on the outcome. (They might be told that outcomes would be determined by coin flips in either case.) See also note 249 *supra*.

There has been empirical work—experiments involving paid student subjects—designed to determine preferences for and beliefs about the adversary system, notably by a group of social psychologists. See, for example, John Thibaut & Laurens Walker, *Procedural Justice: A Psychological Analysis* (1975); Pauline Houlden, Stephen LaTour, Laurens Walker, & John Thibaut, *Preference for Modes of Dispute Resolution as a Function of Process and Decision Control*, 14 J. Experimental Soc. Psych. 13 (1978); Stephen LaTour, Pauline Houlden, Laurens Walker, & John Thibaut, *Procedure: Transnational Perspectives and Preferences*, 86 Yale L. J. 258 (1976); John Thibaut & Laurens Walker, *A Theory of Procedure*, 66 Calif. L. Rev. 541 (1978); Laurens Walker, E. Allan Lind, & John Thibaut, *The Relationship between Procedural and Distributive Justice*, 65 Va. L. Rev. 1401 (1979). The major finding is a preference for adversary procedures and, in particular, for the use of representatives for litigants rather than relying on investigators working for the court or direct interaction between the decision-maker and the litigants. The implications for the present context are uncertain. First, the preference for adversary procedure (which might be viewed as more procedural protection) coincides with less direct involvement of the parties (direct participation often is the least preferred mode in these studies). Second, beliefs about accuracy are no doubt confounded with other values associated with the procedures. (In most experiments, preference rankings correspond with rankings of which procedures are perceived to be more accurate or give relative advantages to the party believed to have the meritorious position.) Mashaw has commented on much of this work by noting that it does distinguish perceptions of means and ends, but “I do not, however, believe that they have been able to isolate perceptions of process that relate necessarily to some dimension of process other than its potential to provide a favorable outcome via either personal participation or the participation of an advocate committed to the claimant’s cause. Indeed, their account of what is at work in a favorable or unfavorable perception of processes seems rather muddled.” Mashaw, *supra* note 195, at 887–88 n.15. Additional difficulties in applying this work include the authors’ failure in some studies to hold information levels (or subjects’ perceptions of these levels) constant; that no valuations are obtained (that is, there is no way to know the weight of the subjects’ preferences; moreover, given the lack of real stakes, subjects’ responses may have been intended to please the experimenter, as Houlden *et al.* observe (at 27) in the *Journal of Experimental Social Psychology*); the inability to know whether the results reflect more than subjects’ stereotypes (as they are not given independent information about the procedures; nor are they involved sufficiently to learn the effects the procedures might have); and the difficulty of determining the source of results (for example, in the study reported in the *Virginia Law Review*, Walker, Lind, and Thibaut argue that participants’ perceived control is central, but in the study they were given virtually no meaningful control; they conclude (at 1417) that the “data suggest that the attorney should facilitate participation by the client in the decisionmaking process,” but this was not the case in their experiment; they do not remind the reader that in their other studies individuals preferred using advocates rather than participating directly). Also, the authors do not attempt to reconcile their work with individuals’ decisions in the world, as in settling lawsuits, choosing alternative methods of dispute resolution, hiring lawyers, and so on.

²⁵⁴ Observe that, in a settlement, both sides forfeit the opportunity to appear personally and participate, implying that settlement destroys value for both parties if participation is indeed valuable to them. (Also, it is typical that settlements are negotiated by lawyers rather than the parties, indicating that the parties attach no value or negative value to their direct involvement in proceedings that affect them—or that the positive value they place on participation is exceeded by the personal benefits they receive from having more expert representation.)

Another is the form of dispute resolution typically specified by contract, and these often are of a simple sort.²⁵⁵ Of particular relevance for *Mathews v. Eldridge*, individuals' private disability contracts presumably do not provide for personal appearances in formal hearings.²⁵⁶ Moreover, in such instances, individuals who agree to summary procedures forgo not only the benefits of greater personal involvement per se but also any positive effect such involvement may have on the accuracy of outcomes. Finally, it is important to recall from Subsection IVA2 that individuals' incentives to promote their interests in claims proceedings, by personal appearance or otherwise, tend to be socially excessive. Thus, even if individuals, at the time disputes arose, did value further participation and were willing to pay for it, satisfying such preferences may be socially undesirable.²⁵⁷

²⁵⁵ See Subsection IIB4a *supra*. For example, in the extensive battle for credit card customers, no company has achieved a large market share by advertising that it offers the opportunity to present disputes in person.

Further evidence is that few individuals take the opportunity to appear in traffic court to dispute violations. Those who do, one suspects, usually face serious consequences (the violation is serious, the last in a series, or will involve a significant increase in insurance rates) or attach a low value to their time (for example, they may be currently unemployed). In addition, the option of disputing in writing may be unavailable. Of course, when the dispute is minor, the process value, if it exists, may be sufficiently small to have little effect in any event.

²⁵⁶ In fact, insurance dispute resolution may provide less by way of formal procedural protection than was present in *Mathews v. Eldridge*.

²⁵⁷ As the analysis in Subsection IVA2 demonstrated, claimants' incentives *ex post*—at the time of adjudication—are socially excessive. This implies that *ex ante*—for example, before knowing whether they would ever become disabled—they would benefit from entering into an arrangement that reduced their prospective rights in adjudication in exchange for lower tax payments to fund the disability program. Evidence from insurance contracts, in which such *ex ante* contracting takes place, is thus particularly probative in this context. See Kaplow, *supra* note 188. Michelman suggests that not all waivers of process rights should be valid: "[I]n this case the contractual commitment is one whereby the one who makes it treats *himself* as a mere means—trading away not just a portion of his future freedom of action but his very claim to be treated as a fully valued person. In that respect the waiver seems distantly akin to voluntary, contractual enslavement." Frank I. Michelman, Formal and Associational Aims in Procedural Due Process, in 18 *Nomos: Due Process* 126, 167 n.66 (J. Roland Pennock & John W. Chapman, eds. 1977). He does not explain, however, how individuals entering ordinary contracts, such as one agreeing to employment as a store clerk or a football player, avoid treatment as less than a fully valued person. (If the distinction does not arise because of waiver, it must be that the infringement on autonomy of, say, not appearing personally at a hearing, is worse in kind than there being no formal decision-making process whatever, as when one's boss orders one to act, no questions asked, at the risk of being fired.) See also Stephen N. Subrin & A. Richard Dykstra, Notice and the Right to Be Heard: The Significance of Old Friends, 9 *Harv. C.R.-C.L. L. Rev.* 449, 458–59, 479 (1974) (authors discuss state action limit to constitutional procedural protection but fail to explain why their rationale applies to government but not private action except by suggesting greater governmental power, implying that abuse of power rather than principles such as human dignity underlie the value of procedural protection, see Subsection 3 and note 268 *infra*).

Alternatively, as with some of the fairness concerns that have been discussed, one might argue that there is a social interest in a procedure that should be considered even if individuals do not value it.²⁵⁸ That is, individuals should be made to pay, or society should be willing to pay, for individuals to win or lose in person. While there are many contexts in which such a procedure would affect accuracy or the implicit burden of proof and others in which it may alter the effectiveness of a result,²⁵⁹ an intrinsic appeal to process suggests that it should be valued even when there are no such effects.²⁶⁰

²⁵⁸ The source and nature of such values is difficult to determine. For example, Mashaw refers to individuals' dignity, see *supra* note 194, at 49–52, but one ordinarily presumes that individuals care about their dignity. In fact, Mashaw suggests that lack of personal participation causes alienation. *Id.* at 50. Yet, without more careful argument and evidence, it is difficult to know whether the alienation comes primarily from losing or from not appearing. (Are those who win similarly alienated? See note 249 *supra*.) One does suspect that individuals will be more alienated when they feel action to be unjustified. Thus, if they were not allowed to present their case effectively or learn of the evidence on the other side that formed the basis for an adverse decision, they may conclude that an outcome is illegitimate. It is not clear, however, that such concerns call for live testimony. An opportunity to present evidence in some form and to be informed of contrary evidence is usually desirable to enhance accuracy in any event. Compare Friendly, *supra* note 194, at 1279–95 (discussing and ranking different elements of a fair “hearing”).

Mashaw also suggests that benefit denials in the disability context involve an adverse moral judgment (that one is not socially excused from supporting oneself, as one claimed). See Mashaw, *supra*, at 51–52. This suggests that an adverse decision involves a sanction in addition to not receiving benefits; if so, this would justify greater accuracy, and then personal appearance would be appropriate if it were a sufficiently useful way of enhancing accuracy. Mashaw argues that personal appearance is inherently necessary because of the moral component of the decision, but the benefits he emphasizes are allowing more emphasis on “a particular claimant’s vocational characteristics, his unique response to his medical condition, and the ultimate predictive judgment of whether the claimant should be able to work.” *Id.* at 52. But whether personal appearance enhances this significantly (beyond the claimant’s written submissions, meeting with doctors and others, and so on) seems more an empirical question. (For example, it is not explained why “a particular claimant’s vocational characteristics” is either moral—in a way that his medical condition, which Mashaw contrasts, is not—or something best observed in person.) In subsequent work, Mashaw de-emphasizes this moral element, at least with regard to whether it is desirable to emphasize it in the decision-making process itself. See Mashaw, *supra* note 87, at 94–95.

²⁵⁹ Being fined for a traffic violation in person, by a judge in robes, may have more of an effect on future behavior than when a violator mails a check. Or denying benefits to an applicant in person may allow for a live, and thereby more effective, explanation of the reasons, which may lead to a better informed decision about whether to appeal or to reapply in the future.

²⁶⁰ The discussion is explicitly addressed to process values in appearance. These seem closely related to values of participation. One would similarly ask why one wishes to participate or why society should value this. If it is to offer information, it is subsumed in the discussion in the text. If it is to interact with the decision-maker, it must be either that one believes that this will improve the reliability of the outcome (which would be subsumed entirely in the preceding discussion) or that it improves the participant in some manner, which raises questions analogous to those discussed in the text.

2. Appearances and Legitimacy²⁶¹

Perhaps a system that allows additional procedures appears to be more “legitimate.”²⁶² Consider first the possibility that the system would be seen as more legitimate because it would appear to be more accurate. If the procedures indeed make the system more accurate, this suggests that accuracy is more valuable than economic analysis typically assumes. The idea may be that compliance with legal commands is enhanced not only by more precisely tailored incentives but also by a belief that the system functions well.²⁶³

A second possibility is that additional procedures would make the system appear more legitimate even though it is not assumed that the procedures make it more accurate. Again, there is the problem from Subsection 1 that few if any individuals actually focus carefully on a hypothetical situation in which a new procedure is added and it has no effect on the outcome. If such a procedure were, nonetheless, seen as contributing to legitimacy, one would wish to know why. One possibility is that a procedure enhances legitimacy because it furthers the sorts of values addressed in Subsection 1.²⁶⁴ Another is that such a procedure

²⁶¹ A different sort of process value is associated with the range of issues considered by decision-makers. Thus, some have suggested that irrebuttable presumptions or other rules that limit debate in adjudication have the undesirable effect of closing off debate in the polity as a whole. See, for example, Laurence H. Tribe, *Structural Due Process*, 10 Harv. C.R.-C.L. L. Rev. 269 (1975). One might ground such a benefit in an instrumental argument, that in some circumstances such a process will improve the quality of a later, ultimate decision. Tribe argues, however, that such a process is not merely a means to enhancing outcomes or legitimacy but a constitutive element of legitimacy. See *id.* at 306–7 & n.114.

²⁶² This subsection does not explore what legitimacy means or why it might be valuable. Of course, given the resulting ambiguity of the subject, one is unavoidably more uncertain about the relevance of any analysis of it.

²⁶³ If the procedures do not produce more accuracy, but citizens mistakenly think that they do, there arises a familiar problem in governance that there is no point in attempting to illuminate here. (As an analogy, one might ask whether the government should adopt a highway plan that results in more loss of life because most citizens mistakenly believe otherwise.) One suspects that, with regard to designing the legal system, the public is indeed poorly informed about how accurate current procedures are or whether various reforms would improve or worsen the situation. It could hardly be otherwise, for the problem is complex and there is virtually no empirical information on the subject available even to experts.

²⁶⁴ Mashaw suggests that legitimacy is needed, which requires “political processes that respond to a democratic morality’s demand for participation in decisions affecting individual and group interests.” Mashaw, *supra* note 194, at 49–50; see *id.* at 45. Interestingly, he refers to “democratic morality’s” demand rather than the “democratic majority’s” demand, rendering the lack of popular support for the demand immaterial. (At one point, he notes the countermajoritarian purposes of constitutional protections, see *id.* at 58, but he does not indicate the countermajoritarian problem in the context of determining eligibility for social security benefits, for which most of the population is potentially eligible.) What

may be seen as preventing abuse of power, as discussed in Subsection 3 below.²⁶⁵

Finally, if one is concerned with appearances and legitimacy, it is useful to consider whether additional procedures tend to enhance rather than detract from the public's view of the legal system (in contrast to lawyers' or legal academics' view). In recent decades, one has heard more complaints about the litigation explosion, the excessive monetary cost and other burdens associated with particular types of lawsuits, and excessive procedures delaying or denying justice than about a serious shortage of procedures or an inadequacy of opportunities to be involved with lawyers or the legal system.²⁶⁶ Although such views in substantial part involve an unsophisticated understanding of the legal system,²⁶⁷ an approach that privileged them may warrant less accuracy and concern for procedure than indicated by the economic analysis presented here.

the "morality" is that demands participation is unclear. He proceeds to discuss human dignity, addressed previously in note 258. Later in the section, he repeatedly notes how participation would provide decision-makers relevant information, encourage them to focus on different issues, and affect their decisions, see *id.* at 50–52, suggesting a close connection between his argument and an instrumental one concerned with accuracy. (The difference may be that he has in mind different criteria than the Supreme Court that the decision-maker is supposed to be applying accurately. Compare *id.* at 44 (arguing that consistency—treating like cases alike—may be the only possible index of accuracy, which suggests that he has in mind whether adjudicators are applying the same criteria, rather than whether they correctly find the facts—which defines which cases are alike).) See also *id.* at 52–54 (discussing "equality" as a separate value but arguing for additional procedural rights largely on grounds of accuracy or to promote dignitary values, as in his discussion of other values).

²⁶⁵ Yet another possibility is that a procedure may enhance legitimacy in part because it may hide inevitable inaccuracies in adjudication. See, for example, *In re Winship*, 397 U.S. 358, 364 (1970) (beyond reasonable doubt standard valuable so as not to leave community "in doubt whether innocent men are being condemned"); Charles Nesson, Reasonable Doubt and Permissive Inferences: The Value of Complexity, 92 Harv. L. Rev. 1187 (1979); Charles Nesson, The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts, 98 Harv. L. Rev. 1357 (1985); Tribe, *supra* note 140. See also Jonathan J. Heohler & Daniel N. Shaviro, Veridical Verdicts: Increasing Verdict Accuracy through the Use of Overtly Probabilistic Evidence and Methods, 75 Cornell L. Rev. 247 (1990) (surveying much of the debate on use of probabilistic evidence and sharply distinguishing between whether using such evidence increases accuracy, which the authors claim it does, and whether it furthers other policies, which it may not); Note, *supra* note 108, at 1807–14 (exploring appearances argument as a positive explanation of hearsay rules but critiquing such an approach on normative grounds).

²⁶⁶ See, for example, Summers, *supra* note 247, at 39 ("Then, too, lay ignorance of process values is widespread. Consider how often laymen seem ready to dismiss procedural rules as 'mere technicalities' even though these are the very rules which must secure most process values."); *id.* at 42.

²⁶⁷ See, for example, Robert C. Clark, Why So Many Lawyers? Are They Good or Bad? 61 Fordham L. Rev. 275 (1992).

3. Abuse of Power

Additional procedural safeguards are advocated—sometimes implicitly—on the ground that they make abuse of power more difficult.²⁶⁸ Thus, for example, if decision-makers are required to state reasons for their decisions, it may be easier to detect inadequate effort or impermissible grounds for decisions.²⁶⁹ The benefits of enhanced detection are two-fold: decision-makers have a greater incentive to behave properly in the first place if their misbehavior is more likely to be observed; and misbehavior that nonetheless occurs may be easier to correct, through appeals of individual cases, removal of decision-makers, or reforms of the process.²⁷⁰ Concerns such as these directly involve problems of accuracy. When decisions are made haphazardly or on criteria other than those deemed relevant, more errors will occur.

There is a further connection between accuracy and the abuse of power. If a system must be structured so that it provides accurate outcomes, it will tend to be more difficult to manipulate, as by prosecuting individuals based on fabricated charges. An abusive government may have the alternative of enacting substantive rules making illegal the behavior it finds threatening. Some such enactments are directly regulated by constitutional provisions, such as the First Amendment, the prohibitions on *ex post facto* laws and bills of attainder, and limits on the government's power to take property. In addition, enactment of offensive substantive law may be self-limiting because it will be more difficult for a government to hide the purposes of its actions. Thus, it may be no accident that so many constitutional protections concerned with controlling government power involve adjudication and often concern its accuracy.²⁷¹

²⁶⁸ Although the relationship between procedural protections and abuse of power is not usually addressed in depth, it is common for advocates of intrinsic process values to offer references to fears of abuse of power in support. See, for example, Subrin & Dykstra, *supra* note 257, at 456–57, 458–59; Tribe, *supra* note 203, at 666–67, 718.

²⁶⁹ The problem of effort is a general problem of incentives that arises when agents must be relied on to implement mandates specified by others—in this context, by the legislature or executive officials. A more particular worry is that decisions will not merely be occasionally erroneous due to insufficient effort but rather will be infected by prejudice.

²⁷⁰ Redish and Marshall distinguish intrinsic arguments concerning the benefits of revealing reasons for official decisions: “An official’s revelation that he has acted arbitrarily in dealing with an individual, though perhaps helpful in certain senses to the individual, does not, in and of itself, treat that individual with dignity. As long as the individual can do nothing to alter the outcome, the revelation does not alter the morally unacceptable fact of arbitrary governmental treatment.” Redish & Marshall, *supra* note 246, at 490.

²⁷¹ Of course, some constitutional procedural protections may interfere with accuracy, as when credible evidence is excluded on account of Fourth Amendment violations or when

Concerns about abuse of power may, however, raise independent issues. First, one may be more concerned with errors that arise, say, due to racial prejudice than as a result of inevitable imperfection in fact-finding. Second, problems involving abuse of power may involve a more complex dynamic. A procedure—say, delegating a decision entirely to an expert—may normally be both cheap and accurate but, in the wrong hands or in tense times, may go awry.²⁷² These factors may affect an assessment of how valuable accuracy is in a given context or the determination of how accurate various procedures would be in practice,²⁷³ and also complicate the economic analysis.²⁷⁴

Finally, it is useful to observe that concerns for abuse of power may fit more comfortably than others with the legal context in which disputes over accuracy have arisen—constitutional challenges concerning procedural due process.²⁷⁵ Some commentators have remarked that a cost-benefit calculus²⁷⁶ such as that employed in *Mathews v. Eldridge* is not

the Confrontation Clause leads to excluding the only available evidence of a crime. See, for example, *Smith v. Murray*, 477 U.S. 527, 544–46 (1986) (Stevens, J., dissenting). This suggests that maximizing accuracy is not in all instances the best way to control abuse. See note 138 *supra*. The point here is merely that it is often useful toward that end. Moreover, when a procedure that reduces accuracy in a given case is adopted to deter abuse, it may be that overall accuracy is enhanced in the process. (Nonetheless, some procedures are advocated on grounds of fairness or dignity that are not claimed to be instrumentally motivated.)

²⁷² As another example, if a fact-finder is allowed to deviate from scheduled damages to fine-tune its award, this added discretion may provide more room for abuse.

²⁷³ Analyzing the latter requires a positive analysis of government institutions, which differs entirely from the sort of analysis presented here.

²⁷⁴ If inaccuracy is due to prejudice rather than adjudicators' inability to produce accurate assessments even while making the best possible inferences from incomplete information, inaccurate awards would be biased, rather than being correct on average (as was assumed in much of the analysis in Part II). It is desirable to eliminate systematic bias—whether produced intentionally or through misunderstanding—regardless of whether actors *ex ante* have precise information about the actual harm they will cause. (Those favored by anticipated error will be underdeterred, and those disfavored will be overdeterred.) This situation is like the case analyzed in Subsection II.A.4, if one assumes that actors anticipate the error *ex ante*. See also Subsections II.A.1c and II.B.3 and note 9 *supra*.

²⁷⁵ One also might defend constitutional review of procedural protections on separation of powers grounds. Thus, if process is insufficient, an agency may be seen as not implementing Congress' commands. Such an approach would have much in common with the positivist one suggested in many of the Court's opinions, see, for example, *Board of Regents v. Roth*, 408 U.S. 564 (1972), and taken to its most extreme form in Justice Rehnquist's plurality opinion in *Arnett v. Kennedy*, 416 U.S. 134 (1974) (defining property rights as limited by the procedure provided), a position rejected by the Court in *Cleveland Board of Education v. Loudermill*, 470 U.S. 532 (1985).

²⁷⁶ It also is not clear that an analysis emphasizing participation, dignity, or other fre-

obviously suited to implementation by appellate courts rather than by legislative bodies or specialized agencies.²⁷⁷ Moreover, the Due Process Clause is a component of the Bill of Rights, much of which is designed to protect individuals against potential government abuse.²⁷⁸ Thus, it is not necessarily appropriate for the courts' tests of the constitutionality of procedures to mirror an economic analysis designed to determine which procedures are socially best.²⁷⁹

VI. CONCLUSION

The economic analysis of accuracy in adjudication suggests that the effects and value of accuracy depend greatly on the legal setting. Accuracy in the assessment of damages is relevant primarily because of how it affects the precision with which legal rules control behavior. But greater accuracy *ex post*, in adjudication, only improves *ex ante* behavior to the extent that individuals are knowledgeable, at the time they act, about the detail an adjudicator later will learn. An important complication is that the anticipation of greater accuracy in adjudication induces individuals to become more informed before they act. Greater accuracy is valuable if it results in a sufficient improvement in behavior to justify its cost in adjudication as well as the additional costs individuals incur to become informed about the legal consequences of contemplated acts. It is noted, however, that individuals' incentives to present information in adjudication are not closely related to the social value of accuracy; in some instances, there are tendencies for these private incentives to be excessive.

quently expressed concerns is more appropriate. One might believe that courts are experts in such matters as participation in adjudication, but presumably they are similarly expert in the accuracy of adjudication. A distinction could, of course, be maintained if "due process" is interpreted, as a matter of definition, to include concerns for participation but not concerns for accuracy.

²⁷⁷ See, for example, Mashaw, *supra* note 183, at 1439, 1450–52; Tribe, *supra* note 203, at 718. There are also problems with designing procedures through piecemeal adjudication. See, for example, note 209 *supra*.

²⁷⁸ See, for example, *Wolff v. McDonnell*, 418 U.S. 539, 558 (1974) ("[t]he touchstone of due process is protection of the individual against arbitrary action of government"); see Mashaw, *supra* note 195, at 898.

²⁷⁹ It does not follow that courts should have no concern for an economic analysis or even that the inquiry into abuse of power is wholly distinct. Rather, it may be, for example, that courts should limit their attention to concerns for abuse of power and attempt to assure that procedures adopted are sufficient to avoid significant threats, to the extent that this is feasible. In some instances, this may involve requiring more or less procedural protection than would otherwise be prudent; in others it may involve imposing different sorts of protection.

Accuracy in determining liability affects deterrence and the costs of imposing sanctions. With regard to the former, greater accuracy, along with increased enforcement effort and higher sanctions, is a way to enhance deterrence. Increasing accuracy differs in that it allows a given level of deterrence to be achieved while imposing sanctions less often, which is an additional benefit when sanctions are socially costly. Increasing accuracy—which reduces false convictions of the innocent and mistaken acquittals of the guilty—is distinguished from raising the burden of proof—which reduces false convictions but increases mistaken acquittals. The discussion emphasizes that accuracy is one of many instruments of the legal system, and the optimal degree of accuracy cannot be separated from how other aspects of the system are designed.

When adjudication determines future rights and obligations, greater accuracy is valuable directly, as it determines the extent to which future actions conform to legal norms. (Accuracy may continue to have values similar to those in other contexts because the anticipation of pronouncements governing future circumstances will influence prior behavior.) The value of accuracy in this context—which receives the most attention in discussions of procedural due process with regard to government entitlements—is often misunderstood because of a failure to distinguish the purposes and effects of accuracy and of changes in the implicit burden of proof.

While this investigation of accuracy is confined to economic considerations, it is observed that the economic analysis incorporates to some extent many other values thought to be implicated by accuracy, such as the fairness of over- or undercompensating accident victims or the cost of mistakenly sanctioning the innocent. Moreover, the discussion suggests that admitting other values may affect the analysis in surprising ways. For example, placing added weight on a concern for mistakenly sanctioning the innocent may favor a legal strategy that employs higher sanctions.

This preliminary study of such a broad subject is inevitably incomplete. First, because the effects and value of accuracy depend greatly on the context, further investigation is necessary both to explore contexts not considered here and to study more carefully those that are examined in order to determine how generally the analysis applies. Second, the analysis can be applied in a concrete situation only after examining how particular legal rules affect accuracy and how accuracy will affect individuals' behavior. Third, this article is confined to economic considerations—notably, the extent to which behavior and the granting of entitlements conforms to the law and the various public and private costs of adminis-

tering the law. Some regard this perspective as including most or all of what is relevant,²⁸⁰ while others believe that significant weight should be accorded to additional concerns. It is to be hoped that the present inquiry will serve two purposes: to suggest that the effects and value of accuracy should not be taken almost entirely for granted, and to outline some important lines of analysis in a wide range of legal settings.

²⁸⁰ One might think that Jeremy Bentham would be the prototype, although his views on procedure are more complicated than they first appear. See Gerald J. Postema, *The Principle of Utility and the Law of Procedure: Bentham's Theory of Adjudication*, 11 *Ga. L. Rev.* 1393 (1977).